



Universidad
Carlos III de Madrid
www.uc3m.es

ESCUELA POLITECNICA SUPERIOR
Grado en Ingeniería de Sistemas Audiovisuales

TRABAJO DE FIN DE GRADO

**Cambio de relación de aspecto en imágenes
basado en contenido**

Autor: Estefanía Estévez Priego

Tutor: Iván González Díaz

24 de Septiembre de 2014

AGRADECIMIENTOS

La terminación del trabajo de fin de grado supone un gran paso en la evolución académica y personal de un estudiante. En este punto, me detengo a echar la vista atrás para recordar el esfuerzo realizado, los obstáculos superados, los logros conseguidos y, por supuesto, para recordar también a todos aquellos que me han acompañado y apoyado a lo largo de este camino.

En primer lugar, agradezco a mis padres la confianza que siempre han depositado en mí. Sus consejos y su apoyo incondicional me han mantenido en pie incluso en los momentos más difíciles, y su fuerza me ha servido de inspiración.

También mis hermanos y el resto de mi familia han estado conmigo, incluida mi abuela, Juliana, hasta el último momento. Sé que estaría orgullosa.

Asimismo, quiero dar las gracias a mis compañeros y amigos de la Universidad Carlos III: Guillermo Mur, Nicolás Jaremek, Juan Luis Sanz, Patricia Rodríguez, Luis Herranz, Álvaro Sánchez, Carmen Alonso, Pilar de la Cruz, Mónica García, Rocío López, Manu Ventero, Alejandro Martínez, Rubén Gómez, Teresa Bergel y otros muchos. No hubiera sido lo mismo sin ellos. También a Luis Fernández, Andrew Hall, Alberto García, Almudena Díaz, Ana Moliner, Jaime y Mariángeles Labiaga.

Quisiera agradecer a mi tutor, Iván González, la oportunidad de trabajar en un tema tan útil e interesante, además de su ayuda y disposición durante la realización de este proyecto.

Por último, doy las gracias a algunos de los profesores que más me han marcado en estos años de carrera: Abelardo Pardo, Amaya Jiménez, Carlos Delgado Kloos, Eduardo Martínez, Enrique San Millán, Fernando Díaz de María, Fernando P. Cruz, Jerónimo Arenas, Julio Villena, Katrin Achutegui, Luis Antonio Azpicueta, Luis Fernando de Inclán, María Celeste Campo, Raquel Crespo, Rodolfo Cuerno y Vanessa Gómez. Gracias por enseñarme.

RESUMEN

Para redimensionar efectivamente una imagen no basta con atender a sus límites geométricos, sino que hay que tener en cuenta su contenido, ya que la importancia de una imagen reside en su interior, no en los bordes que la delimitan.

En este documento se presenta un algoritmo inteligente, bautizado como Seam Carving, que permite cambiar la relación de aspecto de las imágenes priorizando la información relevante contenida en las mismas, evitando pérdidas innecesarias y deformaciones, tanto para reducciones como expansiones.

Dicho algoritmo se basa en la idea de *seams* o caminos de píxeles consecutivos que van desde un extremo a otro de la imagen, ya sea de izquierda a derecha o de arriba abajo, dependiendo de la dimensión que se requiera transformar en cada momento. En nuestro caso, estos *seams* tienen la peculiaridad de portar la menor cantidad de energía dentro de la imagen. Es decir, se calculan de forma que contengan aquella información que resulta más despreciable o aquella que, por su redundancia con el resto del contenido, resulta menos perceptible.

Así, Seam Carving ofrece mejores prestaciones y resultados frente a métodos convencionales de redimensionamiento o cambio de relación de aspecto, limitados por unas directrices que únicamente consideran la alteración de las proporciones de las imágenes desde un punto de vista estrictamente geométrico, ignorando las consecuencias que ocasiona esto en el contenido de las imágenes.

La técnica desarrollada y evaluada en el presente Trabajo de Fin de Grado, permite aumentar o reducir la escala de una imagen para mejorar una composición, encajar un diseño o cambiar la orientación. Pero la implementación eficaz de este método puede utilizarse también para manipular la información propia del contenido y para eliminar elementos específicos de la imagen sin alterar el resto, por lo que se trata de una herramienta muy útil y versátil en el entorno de procesado de imagen.

INDICE DE CONTENIDOS

AGRADECIMIENTOS	i
RESUMEN.....	iii
INDICE DE FIGURAS.....	vii
INDICE DE TABLAS.....	ix
1. INTRODUCCIÓN.....	1
1.1 CONTEXTO.....	1
1.2 MOTIVACION.....	2
1.3 OBJETIVOS DEL PROYECTO	3
1.4 ESTRUCTURA DE DOCUMENTO	4
2. ESTADO DEL ARTE	6
2.1 LAS TECNOLOGIAS AUDIOVISUALES	6
2.2 EL PROCESADO DIGITAL EN IMAGENES	7
2.2.1 INTRODUCCIÓN AL PROCESADO DIGITAL.....	7
2.2.2 CAMBIOS DE RELACION DE ASPECTO	8
2.2.3 SOLUCIÓN COMERCIAL: SEAM CARVING EN ADOBE PHOTOSHOP	10
2.3 MODELOS DE PERCEPCION VISUAL	12
2.3.1 SALIENCIA.....	12
2.3.2 ALGUNOS MODELOS DE SALIENCIA DE INTERÉS	14
3. DISEÑO	16
3.1 FUNDAMENTO TEÓRICO	16
3.2 FUNCIONES DE ENERGIA.....	18
3.2.1 FUNCION DE ENERGIA POR GRADIENTE	19
3.2.2 FUNCION DE ENERGIA POR SALIENCIA	20
4. IMPLEMENTACIÓN.....	24
4.1 HERRAMIENTAS EMPLEADAS.....	24
4.1.1 MATLAB.....	24
4.1.2 GBVS.....	26
4.2 CÓDIGO DESARROLLADO.....	26
5. EXPERIMENTOS.....	27
5.1 BASE DE DATOS.....	27
5.2 PROTOCOLO DE EVALUACION	30

5.2.1 TEST DE FUNCIONES DE ENERGIA PARA SEAM CARVING	30
5.2.2 TEST DE MÉTODOS DE CAMBIO DE RELACION DE ASPECTO	31
5.3 ANALISIS DE RESULTADOS	32
5.3.1 ESTUDIO DE LAS FUNCIONES DE ENERGÍA	32
5.3.2 SEAM CARVING VS RESAMPLING	44
6. GESTION DEL PROYECTO.....	53
6.1 ORGANIZACION DE TAREAS	53
6.2 PRESUPUESTO	54
6.3 MARCO REGULADOR	55
7. CONCLUSIONES Y LINEAS FUTURAS.....	57
7.1 CONCLUSIONES	57
7.2 LINEAS FUTURAS	58
8. REFERENCIAS Y BIBLIOGRAFIA	59

INDICE DE FIGURAS

2.1. Comparación de relaciones de aspecto.	8
2.2. Paisaje en Photoshop.	10
2.3. Paisaje reducido a la mitad de su dimensión horizontal.	10
2.4. Paisaje reducido a la mitad de su dimensión vertical.	11
2.5. Ejemplo de imagen saliencia.	13
3.1. Representación de algunas deformaciones de procesado.	15
3.2. Ejemplo de seam vertical.	16
3.3. Gradientes.	18
3.4. Esquema descriptivo de la obtención de elementos de saliencia.	22
3.5. Mapas de orientación e intensidad.	22
4.1. Ejemplo de gráfica obtenida con Matlab.	24
4.2. Dimensiones de color.	24
5.1. Imágenes de Exterior 1 – Agua y bosque.	27
5.2. Imágenes de Exterior 1 – Artificial.	28
5.3. Imágenes de Exterior 2 – Industrial y carretera.	28
5.4. Imágenes de Exterior 2 – Cultural y vivienda.	28
5.5. Imágenes de Exterior 2 – Comercial y deportes.	29
5.6. Imágenes de Interior – comidas & cenas y zona de trabajo.	29
5.7. Imagen de montaña original y con ratio 0.5.	32
5.8. Imagen de Exterior 1 original y con ratio 0.7.	33
5.9. Imagen de Exterior 2 original y con ratio 0.7.	35
5.10. Imagen de Exterior 2 original y con ratio 0.5.	35
5.11. Acercamiento de objetos de interés.	36
5.12. Deformación de objetos verticales.	37
5.13. Imagen de deportes en entorno natural.	37
5.14. Imagen de deportes en entorno complejo.	38
5.15. Peor caso de cambio de relación de aspecto.	39
5.16. Cambio de ratio en imagen en movimiento.	39
5.17. Comparación de mecanismos.	40
5.18. Regla de los tercios.	41
5.19. Punto de fuga.	41
5.20. Supresión de elementos.	42
5.21. Comparación Seam Carving con Resampling.	44
5.22. Estudio de elementos verticales.	45
5.23. Imagen ejemplo - leonas.	45

5.24. Comparación Seam Carving con Resampling (II).	46
5.25. Comparación para ratio 0.5.	46
5.26. Comparación Exterior 2 - vivienda.	48
5.27. Deformaciones y supresiones en ambos sistemas.	49
5.28. Efectividad Resampling para complementos verticales.	50
5.29. Comparación para imagen de deportes.	50
6.1. Etapas generales del modelo de desarrollo en cascada.	52

INDICE DE TABLAS

Tabla 5.1. Resultados de categoría Exterior 1.	29
Tabla 5.2. Resultados de categoría Exterior 2.	31
Tabla 5.3 Resultados de categoría Interiores.	35
Tabla 5.4 Medias totales de ambos mecanismos.	39
Tabla 5.5. Resultados de categoría Exterior 1 (II).	40
Tabla 5.6. Resultados de categoría Exterior 2 (II).	44
Tabla 5.7. Resultados de categoría Interiores (II).	46
Tabla 5.8. Medias totales de ambos sistemas.	48
Tabla 6.1: Desglose de tareas.	52
Tabla 6.2: Costes de material.	53
Tabla 6.3: Costes de personal.	53
Tabla 6.4: Costes totales.	53

1. INTRODUCCIÓN

Este capítulo es una introducción al contexto, motivación y objetivos del proyecto que incluye una enumeración de la estructura que lo forma, cuya finalidad es proporcionar al lector una idea básica y amplia del proyecto para facilitar su comprensión durante los siguientes capítulos.

1.1 CONTEXTO

La vista es una de las principales capacidades sensoriales del hombre. Gran parte de la comunicación se basa en respuestas a un estímulo o experiencia visual. Puesto que una imagen es precisamente una representación visual, ésta resulta ser el elemento idóneo para transmitir ideas y conceptos de manera efectiva, con una mayor persistencia en la memoria del individuo y rapidez de entendimiento.

Citando el conocido proverbio chino, *una imagen vale más que mil palabras*, se recalca la cantidad de información que puede estar contenida en una sola imagen, y es que las imágenes constituyen un medio de comunicación en sí mismas además de formar parte de otros muchos (ej. Cine, televisión, periódicos). Por ello, un procesamiento y transmisión eficientes resultan imprescindibles.

Los seres humanos estamos continuamente rodeados de imágenes y las aprovechamos cada día en nuestra cotidianidad. Y cada día, los elementos digitales multimedia cobran mayor auge en los medios y en las redes sociales, así como en los equipos tecnológicos cuyo uso es cada vez más extendido y recurrente. Esto conlleva a que, el hecho de que las imágenes sean lo más fieles y flexibles posible, constituya un requisito de calidad que responde a las necesidades y demandas más vanguardistas tanto de usuarios finales como empresas.

En la actualidad, la tecnología se halla en constante crecimiento, de modo que las diferentes plataformas deben adaptarse a las necesidades y preferencias de los usuarios con la mayor celeridad para mantenerse competitivas en el mercado. Teniendo en cuenta la relevancia e influencia de los elementos multimedia en todos los ámbitos sociales, mantener la información y coherencia de las imágenes para distintas relaciones de aspecto, ya sea la pantalla de un Smartphone o la de un televisor de plasma, ya no es

únicamente un requisito de calidad, es un paso adelante en la expansión y perfeccionamiento del procesado multimedia.

1.2 MOTIVACION

El círculo de diversidad y versatilidad de los dispositivos con pantalla se expande continuamente e impone nuevas demandas respecto a los contenidos audiovisuales digitales. De hecho, los diseñadores deben crear alternativas para los contenidos web y elaborar diferentes plantillas para cada tipo de dispositivo en el que vayan a trabajar. Estándares como HTML han optado por desarrollar la capacidad de adaptar el diseño web dinámicamente, al menos en cuanto a texto y estructura se refiere. No obstante, las imágenes suelen mantenerse estáticas, inalterables, sin la posibilidad de actualizar su tamaño de forma automática. Este problema aparece también a la hora de querer encajar una misma imagen en pantallas de distinta relación de aspecto, como se mencionaba anteriormente, o si se desea imprimir a una resolución diferente de la original [1].

Herramientas de redimensionamiento de imagen están presentes en aplicaciones de procesado de imagen, cuya función es cambiar uniformemente las dimensiones de una imagen hasta un tamaño dado. No obstante, el escalado de imagen estándar no es suficiente para esta tarea ya que no es consciente de la información contenida y típicamente sólo puede aplicarse de manera uniforme a ambas dimensiones de la imagen (horizontal y vertical). Las herramientas de recorte están limitadas ya que únicamente permiten eliminar píxeles en la periferia de la imagen y al no prestar atención al contenido resultan ineficientes. No se alcanza el propósito deseado con los medios idóneos.

Por ello, recientemente ha incrementado el interés en el cambio de relación de aspecto en imágenes basado en contenido, donde estas características pueden ser detectadas *top-down* o *bottom-up*. Los métodos *top-down* utilizan herramientas como detectores de caras [2] para localizar zonas importantes de la imagen, mientras que los métodos *bottom-up* dependen de métodos de saliencia visual [3] para construir el mapa de saliencia de la imagen y en base a éste recortar la imagen manteniendo los elementos de mayor relevancia visual.

Por otro lado, la ingente cantidad de datos que atesoran actualmente los contenidos audiovisuales, rodados en alta definición y recientemente también en 3D, necesitan

basarse en una tecnología de procesamiento ricamente evolucionada, además de tecnologías de comunicación y almacenamiento, con las que coexiste.

El procesamiento de la señal, desde la creación de las transiciones más simples y básicas hasta los efectos de fantásticos mundos virtuales debe tener a su disposición las más modernas tecnologías y todos los recursos posibles.

1.3 OBJETIVOS DEL PROYECTO

El objetivo principal del proyecto es desarrollar y estudiar una solución de cambio en la relación de aspecto en imágenes basada en contenido, utilizando para ello un lenguaje de alto nivel, MATLAB, en el que se empleen diferentes alternativas de diseño para el algoritmo *Seam Carving*, y el análisis de su efectividad para distintas categorías de imágenes en función de las alternativas seleccionadas.

Seam Carving utiliza funciones de energía para definir la importancia de los píxeles, y es compatible con varios tipos, tales como el método de magnitud de gradiente, por entropía, saliencia visual, movimiento *eye-gaze*, entre otros. El segundo objetivo es seleccionar dos de estas alternativas para su comparación, además de una herramienta no basada en la premisa de mantener intacta la información relevante para el contenido de las imágenes, con el fin de evaluar la robustez del algoritmo propuesto frente a métodos convencionales.

Como tercer objetivo, mediante la evaluación de los resultados obtenidos en las pruebas comparativas, se pretende buscar y analizar patrones que definan aproximadamente el comportamiento del algoritmo en función de los diferentes tipos de imágenes. Para ello, se recurre a una base de datos categorizada de forma jerárquica en función de su contenido. Concretamente, se desea evaluar la adecuación de estas técnicas a los diferentes tipos de escenas, desde aquellas que representan entornos naturales hasta las que muestran construcciones y elementos generados por el hombre.

Desde nuestro punto de vista, tanto el documento original que presenta Seam Carving como los trabajos posteriores relacionados no han propuesto en ningún momento una evaluación cuantitativa y sistemática de los resultados de la técnica en comparación con otras alternativas. La razón principal se debe a que resulta difícil obtener medidas numéricas sobre la calidad de la imagen resultante que sean capaces de establecer una correlación adecuada con la opinión que un observador tiene de la imagen. En este proyecto, por tanto, se han seguido protocolos de evaluación estandarizados para

calidad en imagen y vídeo, en particular la recomendación ITU-R BT.500-11, a cual propone mecanismos que permiten cuantificar las opiniones subjetivas de los observadores.

Cumplir con este último objetivo requiere a su vez el desarrollo de dos tareas bien definidas: por un lado desarrollar un software que permita la evaluación de los ejemplos por parte de los observadores y, por otro, evaluar numéricamente los resultados para extraer conclusiones sobre el funcionamiento.

1.4 ESTRUCTURA DE DOCUMENTO

El presente documento está dividido en varios capítulos, en los que se desarrollan los conceptos necesarios para comprender el proyecto y proporcionar al lector una descripción general del proceso de diseño llevado a cabo, su implementación y las pruebas realizadas para cambios de relación de aspecto basados en contenido.

El Capítulo 2 está dedicado al ESTADO DEL ARTE de los conceptos y tecnologías relacionados, con el fin de proporcionar suficiente información para conocer el entorno en el que se incluye el proyecto, así como su relevancia en el mismo, y describir un contexto definido que facilite la comprensión del resto del documento.

El Capítulo 3 desarrolla los aspectos principales de DISEÑO incluyendo los fundamentos teóricos en los que se basa la técnica desarrollada, profundizando en los detalles del algoritmo desde un punto de vista técnico y explícito, además de presentar las diferentes alternativas de diseño.

El Capítulo 4 precisa las herramientas y recursos utilizados para la IMPLEMENTACIÓN del software asociado al proyecto, desglosándolos y justificando su elección frente a otras posibilidades.

El Capítulo 5 se centra en los EXPERIMENTOS realizados con sujetos reales, con el propósito de mostrar los resultados de las pruebas de una forma clara, y la interpretación de los mismos mediante un análisis conciso y estandarizado.

El Capítulo 6 desarrolla la planificación del proyecto como conjunto de varias tareas, el presupuesto estimado y el marco regulador en el que se encuentra.

En el Capítulo 7 se enumeran las CONCLUSIONES obtenidas y se proponen posibles mejoras y avances para perfeccionar el proyecto en el futuro.

El Capítulo 8 es una BIBLIOGRAFÍA que enumera de los textos citados a lo largo de todo el documento.

2. ESTADO DEL ARTE

En este capítulo se desarrolla el entorno contemporáneo de la tecnología empleada y bajo estudio desde diversos puntos de vista.

2.1 LAS TECNOLOGIAS AUDIOVISUALES

En nuestra sociedad no surgen los avances tecnológicos trascendentales con mucha frecuencia. Sin embargo, la popularidad de los mismos se extiende a velocidades asombrosas gracias a las redes de la información.

Adaptarse al constante carácter evolutivo de la tecnología supone un esfuerzo continuado, destacando la digitalización del sector audiovisual como una realidad: el ámbito audiovisual ya es digital [4]. Televisión, radio y publicidad están sabiendo obtener el máximo provecho de ello, así como el cine, que a través de su digitalización está consiguiendo crear escenarios nunca imaginados a través de los efectos visuales.

Cada día más exigentes en los efectos transicionales, efectos de croma y mundos virtuales en cine, televisión y otros contenidos visuales, se necesita reservar una gran parte del presupuesto de la producción tras la grabación o toma en bruto. Uno de los cambios más trascendentales fue el paso de la edición lineal a la no lineal, pero la mejora en posibilidades y ahorro de tiempo que ello supuso no hubiera sido posible sin el avance de varias tecnologías de almacenamiento, comunicación y procesado.

Por otro lado, el gran uso y la popularidad de los elementos multimedia en internet son incuestionables, pero la incrustación de componentes audiovisuales en la web no hubiera sido posible sin un incremento del ancho de banda, muy limitado hace unos años, lo que permite una mayor transmisión de información y datos en red.

Hoy en día los usuarios finales tienen la capacidad de utilizar los elementos multimedia al alcance de la mano, ya sean imágenes o vídeos, como una herramienta cotidiana de comunicación, con una accesibilidad más que asequible ya sea a través de un ordenador de sobremesa o un teléfono móvil. Esta situación y los factores sociales que conlleva han impulsado la *necesidad* de compartir contenidos audiovisuales en la red, especialmente a través de las redes sociales y de plataformas multimedia como Flickr, Youtube o Instagram.

2.2 EL PROCESADO DIGITAL EN IMÁGENES

2.2.1 INTRODUCCIÓN AL PROCESADO DIGITAL

El procesamiento digital de imágenes es el conjunto de técnicas que se aplican a las imágenes digitales, generalmente con el objetivo de mejorar la calidad, facilitar la búsqueda de información o preservar su contenido [34]. Se trata de modificar una imagen base para adecuarla a una aplicación específica, ya sea con fines artísticos, tecnológicos, de optimización, etc.

El análisis digital de imágenes tiene muchas ventajas sobre el análisis analógico. Permite el uso de un amplio rango de algoritmos complejos aplicables sobre los datos de entrada y, además de ofrecer respuestas más sofisticadas ante tareas simples, tiene mayores capacidades a la hora de evitar problemas de generación de ruido y distorsión durante el procesamiento.

Las técnicas de procesamiento digital de imagen comenzaron a ser desarrolladas en instalaciones específicas alrededor del año 1960, en el Laboratorio de Propulsión a Chorro, Massachusetts Institute of Technology, de los Laboratorios Bell, de la Universidad de Maryland [37] y algunas otras instalaciones de investigación, con la finalidad inicial de mejorar las imágenes satélite que provenían de la luna. Posteriormente, el interés por este tipo de recurso se extendió hacia ámbitos como la radiología o las telefotografías, que enviaban imágenes por medio del telégrafo o del teléfono, así como hacia otras temáticas más modernas como la renovación de las prestaciones de videoconferencias o el reconocimiento facial.

Con la evolución de las capacidades de los equipos computacionales y la disponibilidad de hardware dedicado en las últimas décadas, los costes de procesamiento se han ido reduciendo considerablemente, y con la velocidad actual de los ordenadores y procesadores de señal, el procesamiento *digital* de imagen se ha alzado de manera indiscutible sobre el resto de tipos de procesamiento de imagen, ya que constituye el método más versátil y también el más económico.

Un caso concreto a mencionar, puesto que hoy en día es un método muy recurrido, es el concepto introducido en 2002 por Raanan Fattal [36] respecto al dominio de gradiente. Según este concepto, resulta más eficiente manipular las diferencias entre píxeles que los valores de los píxeles en sí.

En el mundo cinematográfico, la postproducción de las imágenes está incrementando tanto su éxito que, a modo de ejemplo, en la película Alicia en el país de las Maravillas (2010) dirigida por Tim Burton, fueron necesarios sólo cuarenta días de rodaje pero más de un año de postproducción de contenidos.

2.2.2 CAMBIOS DE RELACION DE ASPECTO

La relación de aspecto de una imagen describe la proporción entre su anchura y su altura. Se calcula como el cociente de la dimensión horizontal entre la dimensión vertical y es representada comúnmente con números enteros con la siguiente nomenclatura X:Y (ej. 4:3) o mediante una notación normalizada (ej. 1.33:1). Generalmente, las dimensiones de una pantalla vienen dadas en las especificaciones del dispositivo en función de la longitud de la diagonal, por lo que para calcular su relación de aspecto se puede recurrir a las siguientes expresiones (1) donde r se corresponde con el ratio, d con la diagonal, h con la altura, w con el ancho y A con el área total.

$$r = \frac{w}{h} \quad h = \frac{d}{\sqrt{(r^2 + 1)}} \quad w = \frac{d}{\sqrt{(\frac{1}{r^2} + 1)}} \quad A = \frac{d^2}{(r + \frac{1}{r})} \quad (1)$$

Las relaciones de aspecto más utilizadas en cinematografía son 1.85:1 y 2.39:1 [31], mientras que las dos relaciones de aspecto empleadas en televisión son 4:3 (1.33:1), formato universal de las pantallas tradicionales hasta el siglo XX, inclusive, y 16:9 (1.77:1) formato universal de alta definición y televisión digital, para televisores panorámicos. En cámaras fotográficas se continúa recurriendo a la relación de aspecto 4:3 junto con la 3:2 y recientemente también la 16:9, pero en fotografía se utilizan además otras como 5:3, 5:4 y 1:1 (imágenes cuadradas).

La conversión de formatos de proporciones diferentes se consigue a través de diferentes métodos. Una opción es aplicar **zoom digital**, con el que la imagen original se agranda hasta que encaja en el área dispuesta en el dispositivo de destino, lo que estrecha el ángulo de visión de la imagen y la recorta. Otro método consiste en añadir **dos franjas negras** para ocupar la totalidad del espacio sin tener que modificar la relación de aspecto de la imagen base. Si se trata de franjas horizontales se denomina *letterboxing* y si se trata de franjas verticales *pillarboxing*. También se puede utilizar un escalado en ambas direcciones, vertical y horizontal, que consigue encajar la imagen a cambio de distorsionarla. En la búsqueda de técnicas óptimas para una buena relación de visionado, Seam Carving se presenta como una solución de peso, ya que evita tanto las deformaciones de la imagen como el desaprovechamiento de espacio en el dispositivo de reproducción.

El ratio 4:3 es un estándar de televisión que ha sido utilizado desde la invención de las cámaras de grabación para cine aunque se emplea también en muchos monitores de ordenador. Fue la relación de aspecto utilizada en las películas mudas de 35mm. A pesar de que sigue utilizándose actualmente en dispositivos como tabletas y algunos monitores, en el siglo XXI este tipo de relación de aspecto para pantallas ha sido desfasado a favor del estándar 16:9 que está presente hoy en día en todas las televisiones de alta definición modernas, además de dominar en el sector de cámaras, monitores y teléfonos móviles. En la figura 2.1 aparece una comparativa ilustrativa de las dimensiones 4:3 y 16:9.



Figura 2.1: Comparación de relaciones de aspecto.

Otra relación de aspecto contemporánea en televisión es la 21:9 [32], para pantallas denominadas ultra panorámicas, cuyo valor de razón exacto es 64:27 y es más próximo al utilizado en cinematografía. Así mismo, este ratio es el que se ha determinado para las nuevas pantallas curvas, se ha empezado a comercializar también para monitores de ordenador y se ha planteado para modelos de teléfonos móviles.

Cabe mencionar por último la relación de aspecto 16:10, puesto que de 2003 a 2008 fue la más común para monitores LCD y ordenadores portátiles, dejando atrás el ratio 4:3. En este caso, el ancho de la pantalla es 1.6 veces su altura, relación aproximada con el número áureo, 1.618. Fue desplazada junto con las demás por el 16:9 y casi relegado al olvido comercialmente. Según la firma de análisis web Net Applications, en octubre de 2012 ya había disminuido en el mercado el número de dispositivos para 16:10 hasta representar menos de un 23%.

Para remarcar la amplia diversidad de las relaciones de aspecto en la actualidad y el abarque de su marco socio-económico, se citan algunos ejemplos: para las tabletas de la marca Apple (iPad) se ha preferido utilizar la proporción clásica 4:3, mientras que tabletas de la familia Nexus y Galaxy Tab usan 16:9. Por otro lado, en las primeras versiones de iPhone se optó por una relación de aspecto 3:2, aunque la gran mayoría de teléfonos móviles, incluidos Android y las versiones más modernas de iPhone, se han decantado también por el ratio 16:9. En marzo de 2014, Samsung se hizo con la patente

para lanzar dispositivos móviles con pantallas 21:9, lo que en muchas industrias visuales se denomina como *Cinema Display*.

En definitiva, parece que las tendencias contemporáneas en cuanto a relación de aspecto son muy variadas y se enfocan hacia modelos de pantallas panorámicas cada vez de mayor tamaño y resolución.

2.2.3 SOLUCIÓN COMERCIAL: SEAM CARVING EN ADOBE PHOTOSHOP

ADOBE PHOTOSHOP¹ es un editor de gráficos rasterizados, es decir, de imágenes en mapa de bits. Actualmente forma parte de la familia Adobe Creative Suite y es desarrollado y comercializado por Adobe Systems Incorporated. Es uno de los programas más populares y utilizados para retoque fotográfico, hasta el punto de ser considerado como el líder mundial del mercado de las aplicaciones de edición de imágenes, siendo así el máximo exponente en cuanto a comercialización de procesamiento de imagen.

En 1987, Thomas Knoll, estudiante de la Universidad de Michigan, escribió un programa que mostraba imágenes en escala de grises en pantallas monocromáticas. Tras trabajar en el desarrollo de esta idea junto con su hermano, en 1988 el programa fue bautizado con su nombre definitivo, Photoshop, y se distribuyeron doscientas copias destinadas a escáneres. Tras una exitosa demostración ante los ingenieros de Apple y Russell Brown, director de arte en Adobe, Photoshop 1.0 fue lanzado en 1990 en exclusiva para Macintosh. [5].

En la versión definitiva de ADOBE PHOTOSHOP CS4², lanzada oficialmente el 23 de septiembre de 2008, se incluía una función de escalado de imágenes basado en contenido que utilizaba el algoritmo Seam Carving y que se ha mantenido hasta las versiones actuales. Esta función cuenta con una opción que permite especificar una relación entre la escala según el contenido y un escalado normal, para poder obtener resultados intermedios. Actúa sobre capas y selecciones, y es compatible con los modos de color RGB, CMYK, Lab y escala de grises así como con todas las profundidades de bits. No obstante, la escala según el contenido no funciona sobre las capas de ajuste, máscaras de capa, canales individuales, objetos inteligentes, capas 3D, capas de video, varias capas al mismo tiempo o grupos de capas.

Las dos alternativas principales que presenta la función de escalado basado en contenido en PHOTOSHOP son: conservación del contenido visual al cambiar la escala

¹ <http://www.adobe.com/es/products/photoshop.html>

² <http://www.adobe.com/support/downloads/detail.jsp?ftpID=4292>

de la imagen y especificación del contenido a proteger durante el escalado [6]. A continuación, en las figuras 2.3, 2.4 y 2.5, se muestran los resultados para una selección de imagen completa, primero reducida a la mitad de su dimensión horizontal y posteriormente el equivalente en una reducción de su dimensión vertical.



Figura 2.2: Imagen original de paisaje



Figura 2.3: Imagen 2.2 reducida a la mitad en la dimensión horizontal.



Figura 2.4: Imagen 2.2 reducida a la mitad en la dimensión vertical

2.3 MODELOS DE PERCEPCION VISUAL

2.3.1 SALIENCIA

A cada segundo nuestros ojos perciben un torrente de datos visuales. Procesar estos datos en tiempo real supone una tarea abrumadora, sobretodo sin la ayuda de mecanismos inteligentes para reducir la cantidad de datos erróneos o filtrar los datos de menor utilidad. Procesos cognitivos complejos de alto nivel, tales como reconocimiento de objetos o interpretación de escena, dependen de un conjunto de datos que han sido transformados previamente por estos mecanismos para que puedan ser tratados. Algunos de estos mecanismos se basan en la atención visual y en la idea de noción y selección por relevancia.

La atención como mecanismo selectivo es comparada a menudo con el concepto de foco de atención o *spotlight*, ya que pretende priorizar el procesamiento visual en la zona de interés, la zona “iluminada” por el foco, en un rango de unos pocos grados del ángulo visual [8]. Con afán de mejorar la metáfora del foco de luz, se intentó que el área de la zona de interés fuese modificable dependiendo del propósito, convirtiendo la atención en una especie de lente de aumento regulable [9]. Cabe mencionar que ninguna de estas teorías tiene en consideración la forma y la extensión del elemento de interés a la hora de establecer el área de actuación. [10]. Este hecho podría parecer trivial, ya que una creencia popular sustenta que la atención actúa *antes* de que los objetos sean reconocidos, pero evidencias experimentales sugieren que la atención puede estar ligada tanto a objetos como a partes de los objetos o incluso a grupos de objetos [11].

En cualquier sistema computacional físico, los recursos para procesamiento son limitados, lo que lleva a congestiones de tipo “cuello de botella” similares a las de la arquitectura de

von Neumann [12] en máquinas digitales convencionales, según las cuales la velocidad de comunicación entre la memoria y la CPU es más baja que la velocidad a la que puede trabajar esta última, reduciendo el rendimiento del procesador y limitando seriamente la velocidad de procesamiento eficaz. Un ejemplo de esta situación es el sistema visual de los primates, en el cual la cantidad de información que llega al nervio óptico, alrededor de 10^8 bits por segundo, excede la capacidad de su cerebro para procesarla completamente y asimilarla en la conciencia. La estrategia desarrollada por la naturaleza para lidiar con esta saturación consiste en categorizar ciertas partes de la información de entrada como prioritarias e ir desplazando el foco de atención de una localización a otra consecutivamente.

Los seres humanos tampoco procesamos toda la información visual que nos rodea, sólo una pequeña fracción registrada por nuestro sistema visual alcanza niveles de procesamiento que llegan a influenciarnos directamente. Así está demostrado en experimentos basados en el fenómeno perceptual *Change blindness* [13], donde cambios significativos en imágenes pasan totalmente desapercibidos bajo condiciones de visualización normales, a pesar de que los observadores no tienen dificultades en reconocer estos cambios una vez les son señalados.

En resumen, los seres humanos contamos con una estructura anatómica visual que es guiada por nuestra atención hacia aquellas partes de una escena en las que se concentra mayor cantidad de información útil, y este proceso puede ser simulado mediante una aproximación computacional.

Mientras que la atención es un concepto general que engloba todos aquellos factores que influyen a los mecanismos de selección, la saliencia caracteriza intuitivamente algunas partes de la escena, ya sean objetos concretos o regiones, que destacan para el observador por encima del resto de elementos adyacentes que lo rodean. Partiendo de estas definiciones, se establecen los diferentes modelos de atención en base al mecanismo utilizado para obtener la saliencia o *saliency*. Un ejemplo de una imagen de saliencia se muestra en la figura 2.5.



Figura 2.5: Saliencia de una imagen.

El mapa de saliencia es un mapa explícito en dos dimensiones que codifica el grado de saliencia de los elementos y los objetos conspicuos de un entorno visual. [14]. Para elaborar este mapa, se tiene en consideración el hecho de que las propias neuronas motoras del sistema humano compiten entre sí hasta que se selecciona una “ganadora” única asociada a una localización determinada, que corresponde con el objetivo de atención a tratar. Inhibiendo esta localización, se permite que el sistema de atención se centre automáticamente en el siguiente objetivo de mayor saliencia visual.

Es importante matizar que, por tanto, la saliencia no es solamente una propiedad física de un estímulo visual, sino que es la consecuencia de la interacción de varios estímulos con un sistema visual, ya sea biológico o artificial. A modo de ejemplo, una persona daltónica tendrá una experiencia de saliencia visual diferente de la que experimentará una persona con una visión de color normal, aunque ambas estén mirando la misma escena física. No obstante, en general, los principales factores que afectan a la saliencia son equiparables para todas las personas.

Por otro lado, hay que tener en cuenta que algunos modelos se incluyen en más de una categoría, y vamos a centrarnos en aquellos que ya hayan sido implementados a nivel de software y que actualmente están preparados para procesar adecuadamente imágenes digitales arbitrarias y devolver su mapa de saliencia pertinente [15]. Concretamente, en la siguiente sección se presentan los modelos cognitivos y los modelos bayesianos.

2.3.2 ALGUNOS MODELOS DE SALIENCIA DE INTERÉS

Prácticamente todos los modelos de atención se han inspirado directa o indirectamente en conceptos cognitivos. Estos modelos cognitivos tienen la ventaja de que expanden nuestra visión desde los pilares biológicos sobre los que se sostiene la atención visual, y nos permiten entender más profundamente los principios computacionales y mecanismos neuronales tanto de este proceso como de otros procesos complejos que dependen de éste, tales como reconocimiento de objetos.

El modelo de Itti ha sido la base sobre la que se han sustentado otros muchos posteriores y pretende replicar el comportamiento de la atención visual humana para escenas tanto naturales como artificiales. Hay al menos cuatro implementaciones de este modelo: iNVT de Itti [16], una serie de herramientas de Saliencia (STB) creadas por Walther [17], VOCUS de Frintrop [18] y un código de Matlab programado por Harel [19]. Posteriormente se han introducido expansiones y propuestas adicionales.

Los modelos bayesianos son una alternativa que combina probabilísticamente evidencias sensoriales (ej. rasgos del objeto) con conocimientos previos (ej. Contexto de la escena) acorde con el teorema de Bayes, para detectar elementos de interés.

Se han propuesto infraestructuras bayesianas para tareas de búsqueda visual [27]. La saliencia de tipo *bottom-up* deriva de la fórmula $\frac{1}{p(f|f_G)}$ donde f_G representa una característica global que resume la densidad de probabilidad de la presencia de un objeto de interés en la escena, basándose en un análisis de los puntos esenciales de la escena en cuestión. Siguiendo esta misma dirección de razonamiento, una de las soluciones propuestas fue optar por integrar linealmente tres componentes: saliencia *bottom-up*, puntos esenciales y rasgos característicos de los objetos para explicar el movimiento de los ojos de los seres humanos cuando buscan personas en una base de datos con 900 escenas. [35]. De estos modelos derivaron otros muchos atendiendo a las mismas premisas.

El principal beneficio de los modelos bayesianos es su capacidad para aprender de datos dados y su habilidad para unificar diversos factores en un mismo principio [28]. Estos modelos pueden utilizar en su favor la estadística de los datos comunes en escenas naturales y otras características que atraen la atención visual del ser humano para predecir su comportamiento.

3. DISEÑO

En este apartado se desarrollan los aspectos fundamentales de diseño, mostrando los mecanismos seleccionados, además de las diferentes alternativas, y exponiendo los conocimientos teóricos necesarios para la comprensión del trabajo realizado, planteado en el presente documento.

3.1 FUNDAMENTO TEÓRICO

La intención del cambio de relación de aspecto basado en el contenido de las imágenes desde un punto de vista computacional es elaborar un algoritmo que priorice determinados elementos o zonas de la imagen de manera juiciosa y con el mayor nivel de eficiencia posible, tanto desde el lado del consumo de recursos como respecto a tiempo y efectividad.

Dada una función de energía y puesto que el objetivo es reducir las dimensiones de la imagen, la estrategia que primero se manifiesta en nuestra mente podría ser la de eliminar los píxeles en orden ascendente, de menor a mayor energía, de forma consecutiva e independientemente de la posición que ocupen. No obstante, esto destruye la forma rectangular de la imagen, porque hay una alta probabilidad de que no se eliminen el mismo número de píxeles en cada fila o columna, como se puede apreciar en la segunda imagen de la figura 3.1.

Para solventar este problema, se puede establecer que siempre se elimine la misma cantidad de píxeles en cada fila o columna como norma. Sin embargo, esto por sí solo no garantiza en prácticamente ningún caso la efectividad del método, ya que no cumple la premisa de respetar las áreas de mayor interés y crea efectos de zigzag, que corresponde con la tercera imagen de la figura 3.1.



Figura 3.1: Representación de algunas deformaciones por procesado.

La siguiente opción a considerar sería eliminar filas o columnas enteras con la menor cantidad de energía de entre todas las que componen la imagen. Pero aun de esta manera se generan deformaciones prominentes. Otra alternativa para evitar esos efectos adversos es el auto-recorte o *auto-cropping*, que consiste en definir una ventana del tamaño deseado para nueva imagen que contenga las zonas de mayor energía. Pero este sistema es muy limitado y poco versátil.



Figura 3.2: Ejemplo de seam vertical.

Esto conlleva a la búsqueda de un operador menos restrictivo que el auto-cropping o el borrado de filas y columnas, pero que mantenga el contenido de la imagen mejor que simples supresiones de píxeles. Aquí es donde nos encontramos con el método Seam Carving y la definición de seams (literalmente “costuras”) internos [1]. Un seam, como se ha definido previamente, es un camino trazado por píxeles adyacentes que van de un lado a otro de la imagen, ya sea de arriba abajo o de izquierda a derecha. Un ejemplo se representa en la figura 3.2.

De forma similar a la eliminación de una fila o columna, eliminar los píxeles de un seam sólo tiene consecuencias locales, en el sentido de que todos los píxeles de la imagen se desplazan hacia la izquierda o hacia arriba, según sea un seam vertical o uno horizontal para compensar los huecos que quedarían “vacíos” en la imagen. Así, el impacto visual es únicamente perceptible a lo largo del camino trazado por el seam, dejando totalmente intacto el resto de la imagen.

Siendo I una imagen de dimensiones $n \times m$, e una función de energía, como se define en la sección 3.2, y considerando un seam vertical, podemos establecer el coste de un seam como:

$$E(S) = E(I_s) = \sum_{i=1}^n \min_s e(I(s_i)) \quad (2)$$

Según esto, el seam óptimo (s^*) que minimiza este coste sería:

$$s^* = \min_s \sum_{i=1}^n e(I(s_i)) \quad (3)$$

Para localizar el seam de mínimo coste en una imagen mediante programación dinámica, centrándonos en el supuesto de una reducción horizontal (seams verticales), el primer paso es recorrer la imagen seleccionada desde la segunda fila hasta la última y calcular la energía acumulada *mínima* M para cada posible camino de píxeles.

Especificamos las coordenadas de cada seam en la imagen como (i,j) y definimos la energía acumulada mínima a través del siguiente algoritmo:

$$M(i,j) = e(i,j) + \min(M(i-1,j-1), M(i-1,j), M(i-1,j+1)) \quad (4)$$

Al final de este proceso, el valor mínimo de la última fila en la matriz M indica el final del seam vertical mínimo conectado. Por tanto, el segundo paso del algoritmo consiste en recorrer inversamente el camino trazado para ese seam; desde el mínimo píxel de la última fila hasta la primera fila, siguiendo el camino óptimo. Para una reducción vertical (seams horizontales) el procedimiento sería el mismo pero intercambiando las dimensiones.

Por estos motivos Seam Carving es un buen candidato para lograr el mejor balance entre la demanda de preservación de energía y la coherencia visual.

3.2 FUNCIONES DE ENERGIA

Si calculamos la energía media de todos los píxeles de la imagen inicial, aunque se eliminen píxeles aleatorios la media energética teóricamente debe mantenerse constante. Sin embargo, al utilizar un cambio de relación de aspecto basado en contenido, esta media aumenta puesto que se están eliminando los tramos (seams) de menor energía, manteniendo los de energía elevada.

Cabe mencionar que hoy en día no existe una única función de energía que se adecúe perfectamente a todos los tipos de imágenes, aunque en general se adaptan adecuadamente en los mismos rangos de redimensionamiento. Las variaciones más notables entre unas y otras se aprecian en el porcentaje de redimensionamiento que son capaces de soportar sin generar deformaciones visuales y en las zonas de la imagen a las que afectan. Shai Avidan y Ariel Shamir en su estudio [1] se decantaron por la función de energía por gradiente como el módulo de la suma de los gradientes y por el histograma de gradientes orientados.

3.2.1 FUNCION DE ENERGIA POR GRADIENTE

Un gradiente de una imagen se define como un cambio direccional de intensidad o color en una imagen. Estos gradientes pueden ser utilizados para obtener información de las imágenes.

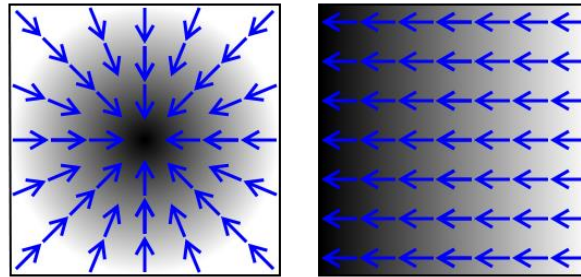


Figura 3.3: Gradientes.

Matemáticamente, el gradiente de una función con dos variables (en este caso la función de intensidad) en cada punto de la imagen es un vector de dos dimensiones cuyos componentes vienen dados por las derivadas de la dirección vertical y de la dirección horizontal. En cada punto de la imagen, el vector de gradientes apunta en la dirección del incremento de intensidad más alto, y la longitud del vector se corresponde con el ratio de variación en esa dirección. Las distancias en las direcciones x e y , en términos de número de píxeles entre dos puntos, se corresponden con dx y dy respectivamente. La figura 3.3 ejemplifica dos tipos de gradientes.

El gradiente de una imagen f viene dado por la siguiente expresión:

$$\nabla f = \frac{df}{dx} \hat{x} + \frac{df}{dy} \hat{y} \quad (5)$$

Donde $\frac{df}{dx}$ es el gradiente en la dirección x , y $\frac{df}{dy}$ es el gradiente en la dirección y .

Para poder trabajar con la imagen a través de MATLAB es necesario procesarla como una matriz y convertir sus parámetros en variables de tipo *double*, ya que es conveniente

mantener la mayor precisión tras muchos cálculos iterativos. Después se calcula la imagen equivalente I en escala de grises y mediante el uso de gradientes se calcula la función de energía según la siguiente fórmula:

$$e(I) = \left| \frac{d}{dx} I \right| + \left| \frac{d}{dy} I \right| \quad (6)$$

Matemáticamente, una norma es³ el tamaño total o magnitud de todos los vectores de un espacio vectorial o matriz. Por simplicidad, podría decirse que cuanto más alta es la norma mayor es el valor contenido en la matriz o vector. Estas normas no tienen una fórmula única y cuentan con diferentes nomenclaturas.

Una evaluación de la norma L-1 y la norma L-2 o euclídea [1] concluye que los resultados proporcionados son bastante similares para ambas en este escenario. En este proyecto se está utilizando la norma L1, conocida comúnmente con el sobrenombre de *Manhattan* norm. [33] Para la diferencia entre dos vectores o matrices, entre científicos computacionales se denomina SAD (Sum of Absolute Difference). A modo informativo, la norma L-2 es utilizada como un estándar de cuantificación para medir diferencias vectoriales.

3.2.2 FUNCION DE ENERGIA POR SALIENCIA

Para obtener la función de energía de una imagen mediante este método se necesita calcular lo que se conoce como mapa de saliencia, lo que requiere una serie de cálculos computacionales sobre la imagen de entrada. Se ha optado por un modelo de percepción visual cognitivo, el modelo Itti et al. [16].

³ <http://rorasa.wordpress.com/2012/05/13/l0-norm-l1-norm-l2-norm-l-infinity-norm/>

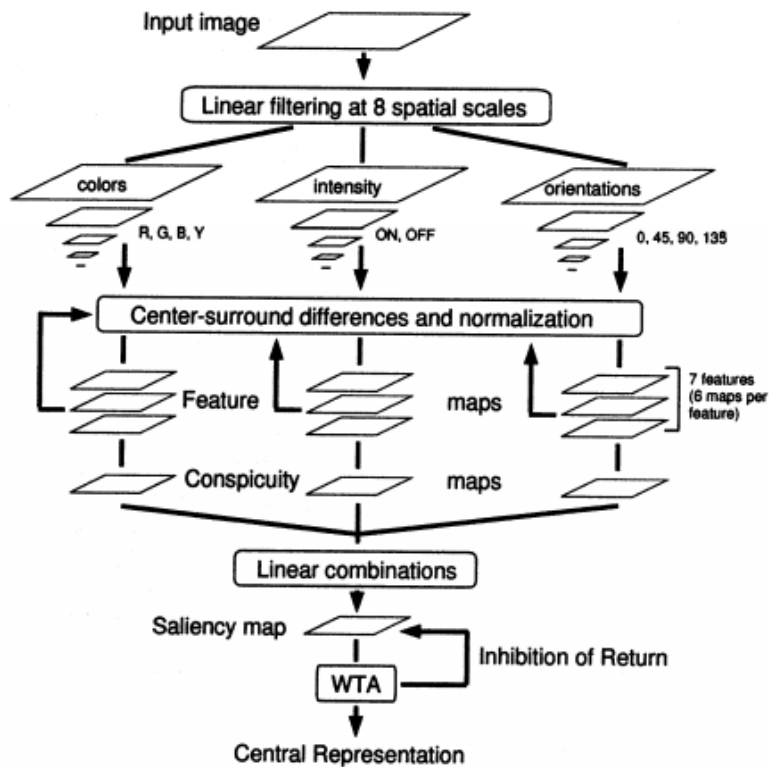


Figura 3.4: Esquema descriptivo de la obtención de elementos de saliencia adaptado de Itti et al (1998) - Obtenida de “Interactions of visual attention and object recognition: computational modeling, algorithms, and psychophysics.”

Este modelo, representado en la Figura 3.4, se basa en tres propiedades o estímulos visuales de bajo nivel: color, intensidad y orientación, y simula el comportamiento de células receptoras del tipo *center-surround*. Nos detenemos un momento para definir el término *center-surround*, que hace referencia a las interacciones entre la zona central y las regiones adyacentes de los campos receptivos de las células fotorreceptoras en la retina. Las células ganglionares de la retina pueden ser excitadas o inhibidas dependiendo del lugar del campo receptivo sobre el cual pase el estímulo visual, dado que los campos receptivos de las células ganglionares son concéntricos. Un campo receptivo de una neurona sensitiva es la región del espacio en la cual la presencia de un estímulo altera la respuesta de dicha neurona. Teniendo estos términos en cuenta, se abren más puertas hacia mejoras de detección de bordes y contrastes en el córtex cerebral, donde tienen lugar la percepción, la imaginación, el pensamiento, el juicio y la decisión.

Dado que este modelo de saliencia es una herramienta utilizada en el proyecto (se ha utilizado un código ya disponible y no se han realizado modificaciones), se ha preferido proporcionar una descripción funcional del mismo, evitando fórmulas y detalles de implementación que podrían dificultar su comprensión.

El proceso de cálculo del mapa de saliencia consta de los siguientes pasos:

1. **Pirámide multi-escala:** Con el objetivo de realizar un procesamiento multi-escala, a partir de una imagen de entrada se genera una pirámide gaussiana de 8 niveles a través de operaciones sucesivas de filtrado paso bajo y diezmado.
2. **Canales de características de bajo nivel:** A continuación, y para cada una de las escalas, se descompone la señal de entrada en varios canales asociados a los 3 estímulos visuales de bajo nivel:
 - a. Color: canales rojo (R), verde (G), azul (B), y amarillo (Y) para dar lugar a dos canales de colores oponentes R-G y B-Y.
 - b. Intensidad: canal de luminancia.
 - c. Orientación: gradientes en 4 orientaciones 0° , 45° , 90° , 135° .
3. **Células *center-surround*:** El siguiente paso es modelar la activación de células del tipo *center-surround*, las cuales se activan ante valores elevados de una característica en la zona central y se inhiben con valores elevados en los alrededores. Para modelar estas células se calculan las diferencias entre mapas en escalas contiguas, simulando el nivel menor (imagen más grande) la zona central, y el nivel mayor (imagen más pequeña) los alrededores. El objetivo de esta operación diferencia *center-surround* es buscar regiones con una característica de valor elevado rodeada de regiones con valores no elevados, lo que se relaciona con una zona de elevada saliencia (una zona diferente de su vecindario).
4. **Combinación de escalas y normalización:** El siguiente paso es, para cada estímulo de bajo nivel, combinar linealmente los mapas obtenidos para cada escala y, posteriormente, normalizar el resultado.
5. **Combinación lineal de estímulos:** En el paso 5 se combinan linealmente los 3 tipos de estímulos considerados (color, intensidad, orientaciones).
6. **Winner-take-all:** El objetivo de esta fase final es, simulando procesos cognitivos de nuestro cerebro, buscar aquellas localizaciones más salientes, las cuales *ganarán* e inhibirán localizaciones cercanas.

Los mapas de los canales de intensidad y orientación procedentes de un estímulo para un conjunto de valores aleatorio, y su posterior normalización, pueden apreciarse en la Figura 3.5 [10].

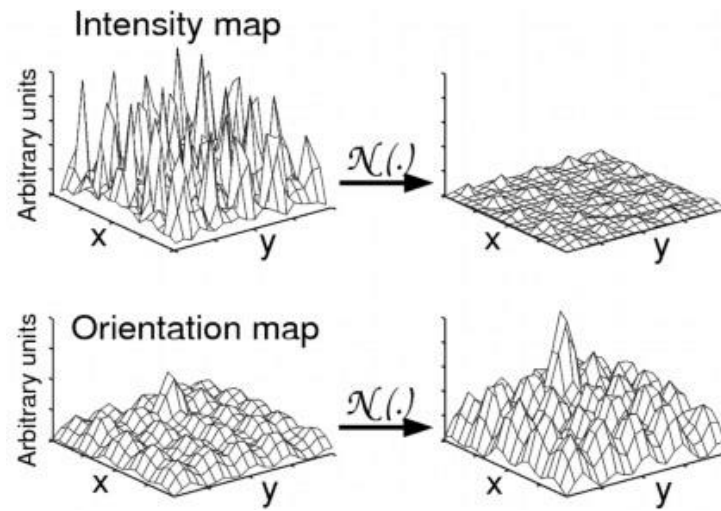


Figura 3.5: Mapas de Orientación e Intensidad.

4. IMPLEMENTACIÓN

En este capítulo se describen las herramientas ya existentes previamente a la realización de este proyecto que han sido utilizadas para su implementación así como los elementos desarrollados específicamente para este mismo fin.

4.1 HERRAMIENTAS EMPLEADAS

4.1.1 MATLAB

Es una herramienta de software matemático que ofrece un entorno de desarrollo integrado y que cuenta con un lenguaje de programación de alto nivel propio. Fue desarrollado por TheMathWorks [16], líder en desarrollo de software para cálculo técnico, y está disponible para las plataformas Unix, Windows, Mac OS X y GNU/Linux.

Nació con el objetivo de tratar de satisfacer las necesidades computacionales de científicos, ingenieros y matemáticos, por lo que el lenguaje que comprende es necesariamente potente e incorpora funciones gráficas más allá de otros lenguajes de programación como C y Fortran.

La ventana principal del entorno de MATLAB cuenta con varias secciones. La más importante es la consola de comandos, que permite interactuar con el programa. Otras secciones son: el directorio de trabajo, que se corresponde con la carpeta virtual con la que se está trabajando; el espacio de trabajo o *workspace*, donde se visualizan las variables ya creadas, y el historial de comandos en el que se almacenan las acciones realizadas previamente en la consola de comandos, para que más adelante puedan ser reutilizadas cómodamente.

Para crear o abrir archivos (.m) en los que se escriben los códigos a ejecutar, se dispone de diversas opciones en la barra de tareas superior de la ventana principal. También incluye otras opciones para simulación y para la creación de interfaces gráficas.

Como su mismo nombre deja intuir, MATLAB (Laboratorio de Matrices) está especializado en las operaciones matriciales, y trata a las variables como tal. Por ejemplo, si se genera una variable $A=1$, MATLAB lo traduce como una matriz de dimensiones 1×1 .

Entre las operaciones disponibles se incluyen la suma, multiplicación, cálculo de la matriz inversa, la matriz traspuesta y el determinante. Permite la selección de componentes, la verificación de dimensiones, el rango y resuelve sistemas de ecuaciones basados en matrices.

Puesto que las imágenes son vistas a nivel de píxel, éstas también son traducidas por MATLAB como matrices, lo que ofrece un amplio abanico de posibilidades computacionales con imágenes en esta plataforma. Además, tanto en dos dimensiones como en tres dimensiones, MATLAB presenta un sistema de gráficos de datos intuitivo con comandos sencillos. Un ejemplo se muestra en la figura 4.1.

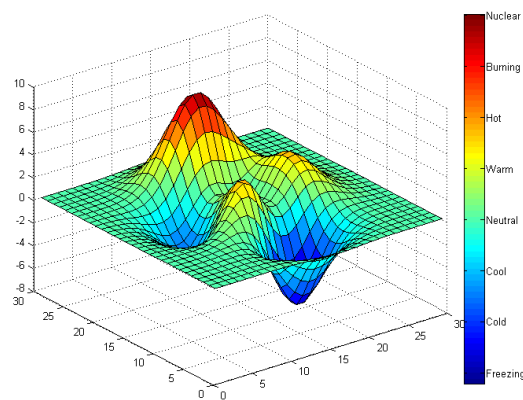


Figura 4.1: Ejemplo de gráfica obtenida con Matlab.

Si la imagen es a color, MATLAB convertirá por defecto los datos de la imagen al espacio de color RGB. Los tres canales de color son representados por la tercera dimensión de la imagen, como se puede ver representado en la figura 4.2.

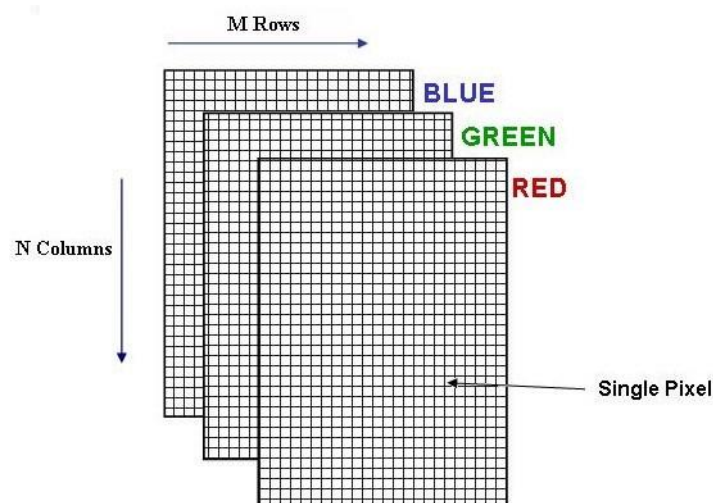


Figura 4.2: dimensiones de color.

El paquete principal de MATLAB en sus últimas versiones incorpora funciones rudimentarias que pueden ser utilizadas para cargar, guardar y realizar funciones personalizadas en imágenes. Pero para la ejecución de operaciones más complicadas la caja de herramientas (*toolbox*) dedicada a procesamiento de imagen permite el uso de procedimientos para: Visualización directa de imágenes en MATLAB, conversiones de espacio de color (ej. Entre RGB, HSV, etc.), agrupamiento de objetos y recolección de datos, análisis de Fourier en imágenes, aritmética de imágenes, operaciones morfológicas, entre otros.

4.1.2 GBVS

Graph-Based Visual Saliency⁴ es un código escrito en el lenguaje de programación MATLAB, desarrollado por el Instituto de Tecnología de California, que permite calcular el mapa de saliencia visual para una imagen o una secuencia de imágenes (video) siguiendo las teorías descritas por J.Harel, C.Koch, y P.Perona, calcular el mapa de saliencia según el estándar Itti, Koch, Niebur o una combinación de ambas. En este proyecto se ha optado por el estándar clásico de Itti y Koch para examinar el comportamiento del algoritmo seam carving basando las imágenes de energía en mapas de saliencia [17].

4.2 CÓDIGO DESARROLLADO

En este proyecto ha sido necesario implementar el algoritmo completo de Seam Carving, ajustándolo a las capacidades de la plataforma descrita en la sección 4.1.1, por ejemplo procurando que el número de bucles fuese el mínimo indispensable, de modo que fuese posible ahorrar costes computacionales y permitir mayor velocidad de procesamiento, aprovechando asimismo los recursos del procesador del ordenador utilizado y los cuatro núcleos de su sistema operativo.

También se requería desarrollar funciones capaces de generar mapas de energía usando gradientes o saliencia, leer y procesar bases de datos manteniendo una estructura predefinida, implementar un software de evaluación de resultados que resultase sencillo e intuitivo para los usuarios que realizasen las pruebas, y un software de cálculos estadísticos automatizado que facilitase la labor de creación de las tablas comparativas en base a los resultados obtenidos.

⁴ www.vision.caltech.edu/~harel/share/gbvs/

5. EXPERIMENTOS

En el capítulo actual se explica el desarrollo en profundidad de los experimentos realizados, cuya finalidad es obtener resultados cuantitativos que puedan ser analizados para extraer conclusiones respecto al comportamiento de los sistemas bajo evaluación y su funcionalidad.

5.1 BASE DE DATOS

El tipo de fotografía influye significativamente en los resultados obtenidos con Seam Carving. Mientras que en fotografías u otras imágenes con predominio de elementos homogéneos (ej. Paisajes) Seam Carving suele ostentar una calidad de resultados excelente, para aquellas imágenes fundamentalmente heterogéneas o en las que hay muchos elementos que requieren ser priorizados (ej. Retratos) los resultados de los métodos de reducción suelen ser considerablemente peores.

El término de género fotográfico designa los diferentes temas tratados en las fotografías. El principal uso de este concepto es disponer la ubicación de las imágenes y facilitar su posterior localización y categorización en los archivos fotográficos y bases de datos.

En la base de datos utilizada en este proyecto se han evitado los retratos en primer plano por su elevada complejidad, porque prácticamente todos los elementos son imprescindibles para el contenido. Por otra parte, se han seleccionado imágenes en su mayoría horizontales, ya que los experimentos que se han realizado se han centrado en mayores disminuciones de la dimensión horizontal que de la vertical, aunque se evalúan todas las posibilidades.

La base de datos jerarquizada sigue concretamente la siguiente estructura:

- ❖ Exterior 1
 - Agua
 - Artificial
 - Bosque/Jungla
 - Montaña/Desierto
- ❖ Exterior 2

- Cultural
- Deporte
- Carretera
- Comercial
- Industrial
- Vivienda

❖ Interior

- Compras y cenas
- Cultural de interior
- Deportes de interior
- Zona de trabajo

Cada una de las catorce subcarpetas mostradas contiene cinco imágenes de la categoría correspondiente, lo que equivale a un total de setenta imágenes. Estas imágenes han sido obtenidas de entre las más de ciento treinta mil expuestas en SUN DATABASE [18] ordenadas en 908 categorías de escena. SUN DATABASE es un proyecto que provee a investigadores de los ámbitos referidos a la visión computacional, la percepción humana, cognitiva, neurociencia, aprendizaje máquina, pruebas de gráficos y robótica, de una colección completa de imágenes anotadas y una amplia variedad de escenas, lugares y diversidad de objetos.

La primera de las categorías se corresponde con Exterior 1 y en ella se recogen imágenes de diferentes tipos de paisajes clasificados según cuatro subcategorías, que pueden separarse fácilmente en dos grupos basándose en la siguiente particularidad: la primera, tercera y cuarta subcategoría se corresponden con entornos naturales en los que no existe ninguna presencia de la mano del hombre (figura 5.1) mientras que la segunda subcategoría contiene entornos naturales en los que se aprecian estructuras artificiales mezcladas con la naturaleza (Figura 5.2).



Figura 5.1: Imágenes de Exterior 1 – Agua y Bosque



Figura 5.2: Imágenes de Exterior 1 - Artificial

La segunda de las categorías es Exterior 2, en la cual se recogen fundamentalmente entornos cotidianos de la sociedad orientados principalmente al aire libre. Predominan las imágenes complejas; entornos industriales, carreteras (Figura 5.3), construcciones de diversos arquetipos, viviendas (figura 5.4), escenas con multiplicidad de objetos y espacios concurridos (figura 5.5). Contiene seis subcategorías.



Figura 5.3: Imágenes Exterior 2 – Industrial y carretera.



Imagen 5.4: Imágenes de Exterior 2 – Cultural y vivienda



Imagen 5.5: Imágenes de Exterior 2 – Comercial y deportes

La tercera de las categorías, desglosada en cuatro subcategorías, son imágenes de Interior centradas en entornos frecuentados por un número considerable de personas, ya sean visitantes de un museo, comensales en un restaurante o trabajadores en una oficina (Figura 5.6). También se incluyen en este nivel imágenes sin presencia de personas, aunque en menor medida, de la misma complejidad.



Figura 5.6: Imágenes de Interior – comidas & cenas y zona de trabajo.

5.2 PROTOCOLO DE EVALUACION

5.2.1 TEST DE FUNCIONES DE ENERGIA PARA SEAM CARVING

Los experimentos han sido evaluados atendiendo a la recomendación ITU-R BT.500-11, en concreto a la sección 5 titulada *The double-stimulus continuous quality-scale (DSCQS) method* [23]. Este método está diseñado para ser especialmente útil cuando no es posible proveer condiciones idóneas para las pruebas que exhiban un nivel completo de calidad. Se trata de un método cíclico en el que se le solicita al evaluador que observe un par de imágenes de la misma fuente. Una de las imágenes debe haber sido sometida

al proceso bajo evaluación mientras que la otra debe ser directamente la imagen original. Lo que debe evaluar la persona encargada de realizar la prueba es la calidad de la procesada con respecto a la original.

Las sesiones no duran más de media hora y se procura fomentar la imparcialidad de las puntuaciones respecto al cansancio visual y evitar el posible tedio que puede suponer este tipo de pruebas. Al final de la sesión se calculan las medias de cada test y se almacenan los datos resultantes para su posterior evaluación y cálculo de las desviaciones típicas.

Para el primer test, los usuarios introducen en un cuadro de texto su valoración atendiendo al siguiente criterio:

- 5: Imperceptible
- 4: Perceptible pero no molesto
- 3: Ligeramente molesto
- 2: Molesto
- 1: Muy molesto

Se ha contado con cinco observadores, cada uno de los cuales debe evaluar todas las categorías de imágenes para cada mecanismo, ya sea con la base de datos de imágenes procesada utilizando **funciones de energía con saliencia o por gradiente**. De esta manera se evita que personas con diferente criterio evalúen diferentes categorías de imágenes, lo que alteraría la validez estadística de los resultados. Estos datos se exponen en las tablas 5.1, 5.2, 5.3 y 5.4.

5.2.2 TEST DE MÉTODOS DE CAMBIO DE RELACION DE ASPECTO

Para el segundo test, el protocolo de evaluación de resultados es un poco distinto al propuesto en la norma ITU-R BT.500-11. En este caso se pretende evaluar las prestaciones de Seam Carving frente a las de un método de redimensionamiento convencional. Para ello los usuarios introducen en un cuadro de texto su valoración atendiendo al siguiente criterio:

- 5: Mucho mejor
- 4: Mejor
- 3: Similar

2: Peor

1: Mucho peor

Los resultados de cada test se guardan en ficheros individuales para poder examinarlos posteriormente y se calculan las medias y las desviaciones estándar siguiendo el mismo procedimiento que para el experimento anterior. Los datos se exponen en las tablas 5.5, 5.6, 5.7 y 5.8.

5.3 ANALISIS DE RESULTADOS

5.3.1 ESTUDIO DE LAS FUNCIONES DE ENERGÍA

En este experimento se han utilizado un total de setenta imágenes para cada test, cada una de ellas se ha presentado junto a otra imagen basada en la original correspondiente pero con una modificación en su relación de aspecto. Concretamente, se ha reducido la dimensión de mayor tamaño al 70% y al 50% de la dimensión original, respectivamente. Se define ratio como la división de la componente de mayor tamaño después y antes del procesado, de modo que en las tablas presentadas a continuación se exponen los resultados para un ratio de 0.7 y un ratio de 0.5 para ambos mecanismos de cálculo de funciones de energía; por gradientes y por saliencia.

Para conocer con detalle el comportamiento de un conjunto de datos, además de las medidas de tendencia central, se requiere conocer también la desviación que presentan los datos en su distribución respecto de la media aritmética, con objeto de modelar la variabilidad en la evaluación por parte de los diferentes observadores. En las tablas se muestran tanto la media aritmética \bar{x} como la desviación típica σ de los resultados obtenidos en cada test, siguiendo la nomenclatura \bar{x}/σ .

5.3.1.1 Subconjunto Exterior 1 - Naturaleza

Exterior 1		Energía por Gradientes		Energía por Saliencia	
	Subcategoría	Ratio 0.7	Ratio 0.5	Ratio 0.7	Ratio 0.5
	Agua	4.6/0.52	3.92/0.58	4.28/0.73	3.8/0.55
	Artificial	3.35/0.82	2.56/0.84	3.68/0.58	2.4/0.85
	Bosque/Jungla	4.52/0.57	3.88/0.58	4.3/0.3	3.5/0.88

	Montaña/Desierto	4.75/0.18	3.84/0.69	4.6/0.4	4.04/0.55
	Media Total:	4.25	3.55	4.21	3.43

Tabla 5.1: Resultados de Exterior 1

Respecto a la tabla 5.1, como puede apreciarse, los resultados obtenidos son particularmente buenos en tres de las cuatro subcategorías. Principalmente, las zonas de montaña son las que cuentan con una media ponderada más alta debido a que, en general, son entornos cuyos elementos principales presentan características muy heterogéneas aunque la composición de la imagen suele carecer de grandes variaciones, de modo que proteger el contenido es mucho más sencillo que para otros casos. Además, las posibles deformaciones son prácticamente imperceptibles gracias a los contornos y texturas poco lineales de los objetos, y especialmente a que no tenemos una imagen preconcebida (a priori) de cómo han de ser estos elementos. Véase la figura 5.7, en la que para una reducción de hasta el 50% a pesar de los cambios producidos en la imagen ésta no resulta molesta. La constancia de las características de cada zona de la imagen hace que estas imágenes sean muy apropiadas para ejemplificar la efectividad del método seam carving con evaluaciones muy buenas y muy baja desviación estándar.

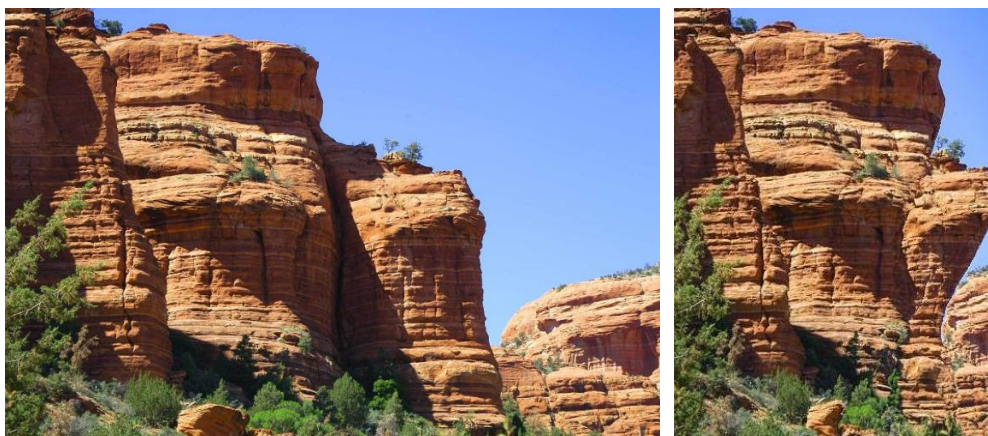


Figura 5.7: Imagen de montaña original y con ratio 0.5

También han dado resultados positivos las escenas de agua, así como las de bosque y jungla, aunque ligeramente más bajos que las citadas anteriormente debido a que estas escenas tienen más diversidad y son ligeramente más complejas. De ahí que la desviación típica sea mayor que en zonas de montaña para un ratio de 0.7. Para ratios menores en las imágenes que presentan errores, dichas deformaciones son caricaturizadas y la desviación estándar aumenta.

En cambio, la subcategoría que se corresponde con paisajes con elementos artificiales sobrepasa por poco el 3 sobre 5 para una reducción del 70%, llegando a ostentar una media de imágenes *molestas* para una reducción del 50%. Esto se debe a que, a pesar de que las zonas de la imagen con predominio natural son procesadas adecuadamente, los objetos artificiales alteran la constancia del conjunto y tienden a deformarse en mayor o menor medida, como ocurre con la imagen 5.8, en la cual a pesar de que el paisaje no presenta deformaciones perceptibles la carretera se ha deformado en una curva imposible. La clave de este comportamiento es que los elementos artificiales sufren restricciones en su geometría y aspecto más duras que en los elementos naturales.



Figura 5.8: Imagen de Exterior 1 original y con ratio 0.7

A su vez, las imágenes procesadas mediante la función de energía por gradiente han conseguido mejores puntuaciones que las imágenes procesadas por saliencia, excepto en la subcategoría “artificial” en la cual este segundo mecanismo protege levemente mejor los objetos artificiales que resaltan en el ambiente natural.

Cabe destacar además que se genera un nivel de molestia mayor en imágenes que han sido sometidas a reducciones de hasta el 50% de su dimensión original que en las que han sido reducidas sólo al 70% de su dimensión original. En imágenes de naturaleza con construcciones artificiales la diferencia es ligeramente más perceptible y se prevé que la notoriedad de estas diferencias se incremente a medida que aumenta el grado reducción.

5.3.1.2 Subconjunto Exterior 2 – Hecho por el hombre

Exterior 2		Energía por Gradientes		Energía por Saliencia	
	Subcategoría	Ratio 0.7	Ratio 0.5	Ratio 0.7	Ratio 0.5
	Carretera	3.65/0.63	2.4/0.99	3.36/0.78	2.36/0.99
	Comercial	2.8/0.54	2.4/0.5	2.96/0.65	1.92/0.52
	Cultural	3.4/1.42	2.68/1.04	3.36/1.78	2.4/1.2
	Deporte	3.35/1.17	3.04/1.3	3.4/1.25	2.72/1.2
	Industrial	3.55/0.92	2.32/1.08	3.16/0.68	2.12/1.11
	Vivienda	3.85/1.07	2.88/0.91	3.76/0.82	2.56/0.93
	Media Total:	3.43	2.62	3.33	2.35

Tabla 5.2: Resultados Exterior 2.

Respecto a la tabla 5.2, se podría deducir a partir de las medias obtenidas que en general los resultados son inadecuados para una visualización cómoda de las imágenes, aunque no indican un nivel de desagrado excesivo. Sin embargo, hay que tener en cuenta que las desviaciones típicas son mayoritariamente altas, indicando tanto discrepancias elevadas entre los observadores como diferencias variables entre imágenes, por lo que es necesario un estudio más profundo de cada caso.

La valoración obtenida en esta categoría, cuyas imágenes se caracterizan por paisajes urbanos, construcciones rurales y otras escenas en las que se muestra la cotidianidad de las personas en ambientes complejos, no completamente naturales, y que cuenta también con escenas protagonizadas por monumentos y otras estructuras artificiales, es una puntuación que refleja considerables molestias visuales. Esto es debido a que hay muchos elementos diferentes, carece de suficientes zonas constantes en la dimensión horizontal y es complicado elaborar una correcta priorización teniendo en cuenta únicamente la energía en función de las diferencias entre píxeles adyacentes o la saliencia propia de los elementos de la imagen, tal como se manifiesta en la figura 5.9.

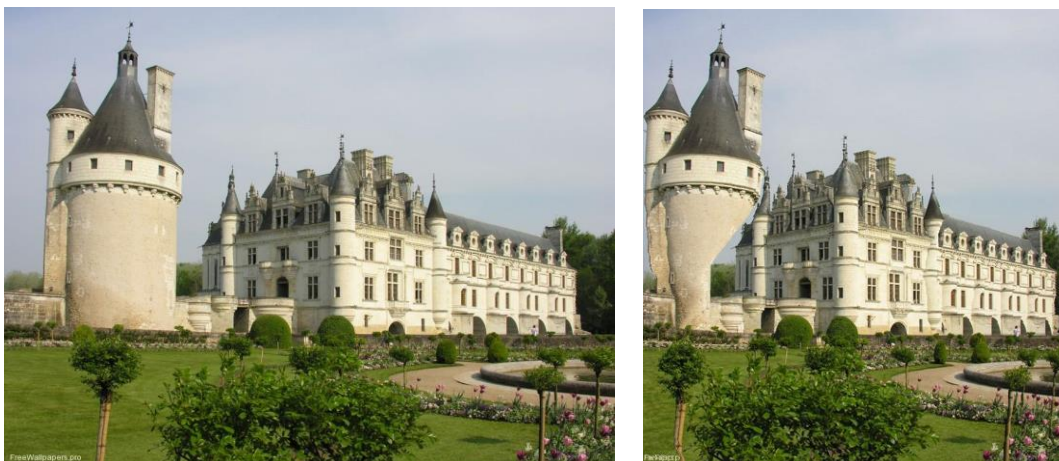


Figura 5.9: Imagen de Exterior 2 original y con ratio 0.7

Estas deformaciones no ocurren en todas las imágenes de esta categoría, pero por estadística tienen más probabilidades de sufrir este tipo de distorsión y que dicha distorsión sea no sólo perceptible sino molesta para la visión humana, especialmente porque reconocemos en seguida los detalles incorrectos al existir fuertes restricciones y conocimiento a priori de las estructuras generadas por el hombre. Por ello, el límite de distorsión es más bajo que en la categoría anterior (Exterior 1) al haber más cantidad de información relevante para el contenido.



Figura 5.10: Imagen de Exterior 2 original y con ratio 0.5

Como era de esperar, los resultados empeoran a medida que el ratio de escala disminuye, esto es, según la imagen queda más reducida, puesto que llegado a cierto límite toda la información es estrictamente relevante para mantener el contenido de la imagen. Este límite es más alto para imágenes de naturaleza sin construcciones artificiales.

En esta categoría, las imágenes reducidas con un ratio 0.5 se clasifican en la mayoría de los casos como inviables, tal como se puede apreciar en la figura 5.10, en la cual la torre

del castillo ha sido prácticamente borrada. El agravamiento de las deformaciones a medida que disminuye el ratio de escala es muy notable.

Las imágenes de Exterior 2 que mejor resultados ofrecen son aquellas cuyos elementos principales están bien definidos y separados entre sí, ya que facilita la labor de localización y diferenciación de los mismos, como se puede ver en la figura 5.11. Además, de esta manera cada uno de los objetos relevantes para el contenido será menos afectado en el proceso de reducción.

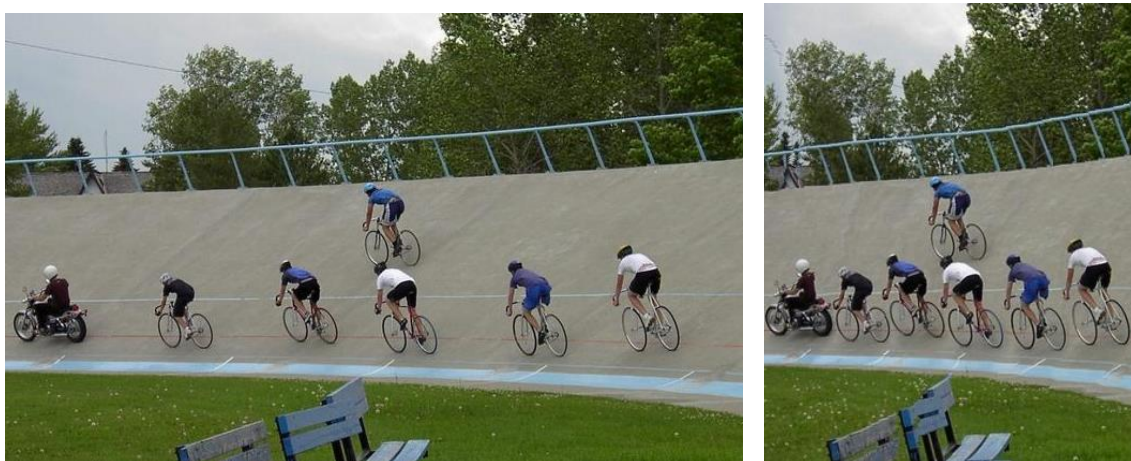


Figura 5.11: Acercamiento de objetos de interés. (izq. Original, der. Ratio 0.5).

También es importante que no aparezcan líneas perpendiculares a la coordenada de reducción. En una reducción horizontal, los seams recorren la imagen de arriba abajo eliminando un píxel de cada fila y por tanto, siempre que haya una línea horizontal ésta se va a ver afectada y se genera la impresión de que esa zona de la imagen ha sido pixelada. El resto de líneas rectas son asimismo muy vulnerables en este tipo de procesado ya que tienden a deformarse.

Los elementos de estructura recta vertical son igualmente vulnerables. Por ejemplo, en el caso concreto de la imagen representada en la Figura 5.12, en la que hay varios árboles de tronco fino entre maleza muy heterogénea, puesto que el algoritmo da prioridad a las partes con mayor energía, es decir, con mayores variaciones entre píxeles contiguos, el método tiende a eliminar parte de los troncos verticales haciendo que se vuelvan demasiado estrechos y se deformen, llevando al usuario a percibir una incongruencia visual.



Figura 5.12: Deformación de objetos verticales. (izq. Original, der. Ratio 0.7)

Se puede concluir, por tanto, que en imágenes con mucha entropía en todas las zonas, elementos relevantes para el contenido pueden sufrir perjuicios si son ligeramente más homogéneos que las otras zonas, aunque éstas tengan menor relevancia visual.

Se destaca, como se ha mencionado anteriormente, que dentro de una misma subcategoría puede darse el caso de que algunas imágenes tras ser procesadas no presenten ningún tipo de molestia, mientras que otras resulten inverosímiles. En la subcategoría “cultural” la desviación típica es de 1.42 ya que hay imágenes con una puntuación media de 1.75 mientras que otras han logrado llegar al 4.75, dependiendo de la simetría, complejidad y elementos que conforman el edificio o lugar cultural de la fotografía. Para la función de energía basada en saliencia esta desviación llega a 1.72, el valor más alto de desviación típica de todas las subcategorías. Un comportamiento similar sucede con la subcategoría “deportes”, ya que influye significativamente el entorno en el que sucede el evento deportivo. No supone la misma complejidad una pista de atletismo que una pista de esquí, como se señala en las figuras 5.13 y 5.14.



Figura 5.13: Imagen de deportes en entorno natural (izq. Original, der. Ratio 0.7).

En la figura 5.13 los esquiadores no han sido perjudicados en absoluto, el algoritmo se ha centrado en las zonas que los rodean sin problema gracias a que se trata de un paisaje

de fondo bastante uniforme. Por otro lado, en la figura 5.14 se observa que la imagen de la derecha, que es aquella que ha sido procesada, ha padecido distorsiones sumamente molestas que han afectado tanto a la pista de atletismo como a los atletas. Cabe destacar nuevamente las restricciones geométricas preconcebidas, en este caso respecto a cómo son las líneas de una pista de atletismo, constituyen una fuente de distorsión importante.



Figura 5.14: Imagen de deportes en entorno complejo (der. Original, izq. Ratio 0.7)

5.3.1.3 Subconjunto Interiores

Interior		Energía por Gradientes		Energía por Saliencia	
	Subcategoría	Ratio 0.7	Ratio 0.5	Ratio 0.7	Ratio 0.5
	Compras y cenas	3.35/1.23	2.4/0.91	3.28/1.1	2.6/0.97
	Cultura de interior	3.2/1.27	2.44/0.96	3.28/1.07	2.4/1.2
	Deportes de interior	3.4/1.63	2.8/1.64	3.28/1.62	2.72/1.5
	Zona de trabajo	2.2/1.1	1.56/0.62	2.12/1.06	1.32/0.46
	Media Total:	3.04	2.3	2.99	2.26

Tabla 5.3: Resultados de la categoría Interior.

La calificación obtenida para imágenes de interior es la peor de todas las categorías. Todos los resultados presentan signos evidentes de molestia, y con desviaciones típicas altas. Los observadores detectan rápidamente las distorsiones. La razón de esta

situación es que en esta categoría hay un alto predominio de personas y en la mayoría de las escenas éstas no se encuentran claramente separadas de otros elementos de interés, lo que conlleva deformaciones muy visibles que enseguida captan la atención del observador. Sucede especialmente en la subcategoría de “zona de trabajo”, que ha conseguido la peor puntuación de todas las subcategorías tal como se puede ver en la figura 5.15. Dichas deformaciones empeoran notablemente al pasar de una reducción de ratio 0.7 a una de ratio 0.5.



Figura 5.15: Peor caso de cambio de relación de aspecto (izq. Original, der. Ratio 0.7)

No obstante, al igual que en la categoría anterior aquellas imágenes cuyos diferentes elementos están bien definidos, es decir, que están separados y sobre fondos más homogéneos, exhiben mejores soluciones para el procesamiento de cambio de relación de aspecto basado en contenido. En la Figura 5.16 se puede advertir que se han respetado los objetos en movimiento y se ha eludido el elemento principal que, en este caso, es el niño de casco rojo. Otros elementos clave como las pelotas de tenis se han mantenido intactos y como consecuencia se han acercado entre sí en la nueva imagen al desaparecer prioritariamente las zonas entre ellos.



Figura 5.16: Cambio de ratio en imagen en movimiento. (izq. Original, der. Ratio 0.7)

Nuevamente, las diferencias entre las imágenes procesadas a través de un mapa de saliencia y aquellas que se han obtenido mediante un mapa de energía por gradientes son prácticamente inapreciables. Los resultados numéricos basados en las valoraciones de los observadores denotan un levísimo empeoramiento en la mayoría de las imágenes de prueba para el mecanismo basado en saliencia. En la figura 5.17 se muestra una misma imagen procesada mediante gradientes (imagen de la derecha) y utilizando el concepto de saliencia (imagen de la izquierda), pero la diferencia es tan sutil que únicamente se percibe si se superponen las imágenes.



Imagen 5.17: Comparación de mecanismos (izq. Saliencia, der. Gradientes).

Aprovechando el ejemplo de la imagen anterior, cabe destacar que de entre todos los empleados el que más daños visuales ha sufrido ha sido aquel que está situado más cerca en la profundidad de campo y que ocupa mayor espacio de la imagen. Las deformaciones que hayan podido sufrir el resto de personas de la fotografía, al ser más pequeñas, menos visibles y menos importantes para el contenido debido a su repetición, son consecuentemente menos perceptibles. Por tanto, la posición y el tamaño de los elementos son también influyentes.

Respecto a la localización de los elementos de interés, en las artes visuales como fotografía o diseño se utiliza la regla de los tercios como una forma de composición para ordenar objetos dentro de la imagen. Esta regla divide una imagen en nueve partes iguales, utilizando dos líneas imaginarias paralelas y equidistantes de forma horizontal y dos más con las mismas características de forma vertical. Si se distribuyen los objetos de la escena sobre las cuatro intersecciones de estas líneas, llamadas puntos focales, además de crear una imagen estéticamente agradable y equilibrada, la atención del observador tenderá a centrarse en dichos objetos, como sucede en la figura 5.18 [24].



Figura 5.18: Regla de los tercios – Obtenida de graficoswebalacarta.com

Del mismo modo, aquellos objetos situados en un punto de fuga recibirán más atención, ya que la vista del ser humano se redirige intuitivamente hacia ellos, y si se producen deformaciones en ubicaciones con estas características éstas serán mucho más perceptibles para los observadores. El punto de fuga es el lugar geométrico en el cual las proyecciones de las rectas paralelas a una dirección dada en el espacio, no paralelas al plano de proyección, convergen. Como ocurre con los raíles de las vías del tren en la imagen 5.19. En este ejemplo los observadores tienden a fijarse antes en la distribución de los árboles del fondo que en los raíles desfigurados.



Figura 5.19: Punto de fuga. (izq. Original, der. Procesada).

Por último, hay que tener en consideración que puede ocurrir que una imagen sea de mejor calidad para una reducción de tamaño mayor. Este es el caso en el que al eliminar un 30% de la imagen algunos elementos son suprimidos sólo en parte y adquieren formas incoherentes, mientras que al reducir hasta un 50% estos elementos desaparecen por completo y de este modo el observador no percibe ninguna molestia ya que el objeto contradictorio ha sido eliminado de la imagen. Este fenómeno puede intuirse en la figura 5.20, en la que para una mayor reducción ambas partes simétricas

de la fachada de este centro comercial tenderían a desaparecer definitivamente. No obstante, este escenario concreto tiene lugar en muy pocas ocasiones.



Figura 5.20: Supresión de elementos. (izq. Original, der. Procesada con ratio 0.7).

5.3.1.4 Resultados globales

Se concluye que para ambos mecanismos Seam Carving se obtienen imágenes de mayor calidad tras el procesamiento si los elementos relevantes contenidos tienen características muy heterogéneas con respecto al fondo, si están bien definidos y separados unos de otros, si no presentan líneas perpendiculares a la coordenada de reducción, si el entorno en el que se ubican es sencillo y la composición general de la imagen es relativamente constante, y si la imagen tiene suficiente resolución para que el algoritmo halle correctamente y con la mayor precisión posible las imágenes de energía.

Las imágenes más complicadas generan mayor variación de resultados respecto a la media y cuantos más seams son eliminados, mayor es la probabilidad de sufrir deformaciones o cambios indeseados en la imagen puesto que, llegado a determinado límite, no quedan zonas cuya supresión pueda ser considerada despreciable.

	Ratio 0.7	Ratio 0.5
Media total de gradientes:	3.57/5 – 7.14/10	2.82/5 – 5.64/10
Media total de saliencia:	3.51/5 – 7.02/10	2.68/5 – 5.36/10

Tabla 5.4: Medias totales de ambos mecanismos.

Para un ratio 0.7 las valoraciones globales para ambos mecanismos rondan los 3.5 puntos sobre 5, lo que equivale a una media de imágenes cuyas deformaciones son en general perceptibles pero suponen muy poca molestia visual. Para un ratio de 0.5 en ambos mecanismos, el resultado global indica que, en media, las imágenes generan molestias relevantes para la comodidad visual.

La clara diferencia de valoraciones entre las diferentes categorías es mucho mayor que la diferencia experimentada entre las distintas subcategorías. Por otra parte, las diferencias presenciadas entre los dos mecanismos bajo estudio son muy reducidas a pesar de la gran diferencia computacional que existe entre ambos. A raíz de estos resultados, y para esta base de datos, **la función de energía de una imagen calculada a través de los gradientes** se consolida como la opción más adecuada.

5.3.2 SEAM CARVING VS RESAMPLING

En este apartado se ha realizado un estudio que compara los resultados obtenidos con el algoritmo Seam Carving, basado en el mecanismo de cálculo de función de energía mediante gradientes, con los resultados obtenidos para las mismas imágenes procesadas mediante una función convencional de redimensionamiento orientada a imágenes, en concreto la función utilizada ha sido *imresize*, la cual está embebida en Matlab.

Se han utilizado un total de setenta imágenes para cada test, recurriendo a la misma base de datos que en el experimento anterior. Cada una de estas setenta imágenes originales se ha presentado junto a la misma imagen procesada mediante Seam Carving y junto a una tercera imagen procesada mediante redimensionamiento convencional. Se han comparado los resultados para reducciones con ratios de 0.7 y 0.5. En cada caso se evalúa la calidad de la imagen procesada con Seam Carving con respecto a la remuestreada siguiendo los criterios presentados en la sección 5.2.2.

5.3.2.1 Subconjunto Exterior 1 - Naturaleza

Exterior 1		Seam Carving vs. Resampling	
	Subcategoría	Ratio 0.7	Ratio 0.5
	Agua	3.45/0.94	4.2/0.74
	Artificial	3.25/1.21	4.1/0.76

	Bosque/Jungla	3.1/1.26	3.84/0.67
	Montaña/Desierto	3.35/1.02	3.7/0.78
	Media Total:	3.29	3.69

Tabla 5.5: Resultados de Exterior 1 (II).

Examinando los datos obtenidos en la tabla 5.5 para un ratio de 0.7, medias que rondan el 3 sobre 5 y variaciones de más de un punto respecto a la media, se puede intuir que independientemente de la subcategoría en cuestión, aunque hay imágenes para las que Seam Carving consigue unos resultados muy buenos, en otras cuyo procesamiento basado en contenido no resulta tan satisfactorio puede ser preferible para el usuario que las proporciones de los elementos estén desfiguradas a que falten partes cruciales en la composición. No obstante, resulta necesario analizar el origen de estos resultados en profundidad.



Figura 5.21: Comparación Seam Carving vs Resampling
(izquierda - Seam Carving, centro - Original, derecha - Resampling)

En la figura 5.21 se exhiben de izquierda a derecha: la imagen procesada con Seam Carving, la imagen original y la imagen procesada utilizando Resampling, todas para un ratio de 0.7. La ventaja de Resampling en este caso es que no ha recortado ninguna parte de los delfines, que son los elementos clave de la imagen. No obstante, es evidente que se ha producido un alargamiento antinatural, mientras que en Seam Carving se han mantenido perfectamente sus proporciones físicas.



Figura 5.22: Estudio de elementos verticales (Izqda. Seam Carving, der. Resampling)

En la figura 5.22 se ve claramente que en la imagen de la izquierda, procesada con Seam Carving, los troncos de los árboles se han estrechado hasta tener el mismo grosor que sus ramas. Por otro lado, en la imagen de la derecha, procesada con Resampling, al tratarse de elementos principalmente verticales apenas se nota que hayan sufrido un alargamiento. Para este tipo de imágenes, con predominio de objetos paralelos a la coordenada de reducción, Resampling es mucho más eficaz.

Prestando atención a la figura 5.23, que muestra una imagen obtenida directamente de la base de datos sin procesar, y comparándola con las imágenes resultantes de la figura 5.24, Seam Carving a la izquierda y Resampling a la derecha, se puede apreciar una situación similar a la comentada anteriormente sobre la figura 5.21.



Figura 5.23: Imagen ejemplo - leonas.



Figura 5.24: Comparación de Seam Carving vs Resampling II

Ambos sistemas han modificado el contenido de la imagen pero de formas diferentes. Seam Carving ha protegido a las leonas como objetos principales del conjunto, sin alterar sus proporciones a cambio de sacrificar el contorno de la roca. Resampling ha deformado linealmente todos los elementos, motivo por el cual la imagen no resulta molesta a simple vista, aunque las proporciones y forma de los cuerpos de las leonas han sufrido cambios significativos. En este punto, basándonos estrictamente en salvaguardar el contenido, Seam Carving ha realizado su función con eficiencia, puesto que se aprecia el tamaño, la forma y la postura de los animales tal como en la imagen original. No obstante, a nivel de comodidad visual la valoración es totalmente subjetiva.



Figura 5.25: Comparación para un ratio de 0.5

Cuando el grado de reducción aumenta, para un ratio de 0.5, los observadores se decantan más por las imágenes tratadas con Seam Carving a pesar de su notable empeoramiento porque las deformaciones creadas por el sistema de Resampling se vuelven excesivamente antinaturales. En la figura 5.25 se percibe claramente que la

persona que se encuentra en el puente ha sufrido desperfectos durante el procesado en ambas imágenes, y sin embargo la calidad del resto de la imagen y la verosimilitud de la composición es preferible en la imagen de la izquierda, que se corresponde con el procesado mediante Seam Carving, frente a la imagen de la derecha, que se corresponde con el procesado mediante Resampling.

5.3.2.2 Subconjunto Exterior 2 – Hecho por el hombre

Exterior 2		Seam Carving vs. Resampling	
	Subcategoría	Ratio 0.7	Ratio 0.5
	Carretera	2.75/0.35	3.4/1.09
	Comercial	2.8/0.94	3.7/0.78
	Cultural	3.1/1	3.35/1.04
	Deporte	3.8/1.2	3.75/1.15
	Industrial	3.05/1.04	3.45/0.92
	Vivienda	3.5/0.86	3.85/0.8
	Media Total:	3.16	3.58

Tabla 5.6: Resultados Exterior 2 (II).

Los datos recogidos para esta categoría continúan la misma línea que los de la categoría anterior. Seam Carving, en media, es considerado mejor que el método convencional elegido para cambios de relación de aspecto en imágenes, pero no con resultados suficientemente concluyentes. Hay muchas condiciones y variantes que influyen en estos datos finales, como se puede entrever por las altas desviaciones estándar que mayoritariamente indican variaciones de alrededor de un punto respecto a la media aritmética.



Figura 5.26: Comparación Exterior 2 - Vivienda

La figura 5.26 representa de izquierda a derecha: La imagen procesada mediante Seam Carving, la imagen original y la imagen procesada utilizando Resampling. En este ejemplo Seam Carving ha mantenido efectivamente las dimensiones de la fachada de la imagen aunque sacrificando la profundidad de la casa. Por otro lado, Resampling ha afectado por igual a todas las partes de la imagen sin discriminación, lo que ha ocasionado una deformación mucho más perceptible que usando Seam Carving. Se puede generalizar por tanto, que en imágenes cuyos elementos de mayor relevancia visual estén bien caracterizados y definidos, Seam Carving genera resultados más acordes con la realidad que Resampling.

En esta categoría, se han logrado resultados suficientemente satisfactorios para Seam Carving, especialmente para las escenas de vivienda y deporte, aunque más bajos que en la categoría anterior, lo que quiere decir que los resultados de Seam Carving frente a Resampling, aunque buenos, no son tan decisivos.

Es importante resaltar que en este experimento se comparan entre sí ambos sistemas de redimensionamiento, no la calidad en sí de las imágenes. En la figura 5.27 se observa que, frente a una desfiguración que abarca a todos los componentes de la imagen pero que no deriva en una incomodidad visual grave, si se procesa con Seam Carving las personas que aparecen en el plano principal es casi suprimida. Ninguno de los dos métodos ha funcionado con la calidad suficiente, por lo que ambos obtienen una puntuación similar por parte del observador.



Imagen 5.27: Deformaciones y supresiones indeseadas en ambos sistemas.

Paralelamente, existen muy pocas variaciones entre la comparación para imágenes procesadas para un ratio de 0.7 e imágenes procesadas para un ratio de 0.5, excepto en las subcategorías “carretera” y “comercial” en las que la balanza se inclina un poco más hacia Seam Carving con un aumento en el promedio de más de 0.5 puntos. En las demás subcategorías la balanza se mantiene estable ya que cuando se producen efectos adversos en las imágenes procesadas dichos efectos incrementan de forma equiparable en ambos casos según aumenta el porcentaje de reducción, es decir, según disminuye el ratio.

5.3.2.3 Subconjunto Interior

Interior		Seam Carving vs Resampling	
	Subcategoría	Ratio 0.7	Ratio 0.5
	Compras y cenas	3.2/1.28	3.7/1.08
	Cultura de interior	3.1/0.96	3.6/0.84
	Deportes de interior	3.55/1.76	3.9/1.2
	Zona de trabajo	2.3/0.91	3.1/1.09
	Media Total:	3.03	3.56

Tabla 5.7: Resultados de la categoría Interior.

Una vez más, el conjunto de imágenes de Interiores recoge las valoraciones más bajas, aunque éstas aumentan para mayores reducciones debido a que en muchas de las imágenes Resampling empeora más rápidamente conforme disminuye el ratio que Seam Carving. No obstante, para bajas reducciones, en esta categoría en la que predomina la presencia de personas y elementos complejos, Resampling salva con más eficiencia los elementos finos y composiciones con numerosos objetos, como ocurre en la imagen 5.28 con las piernas de los visitantes del museo.



Figura 5.28: Efectividad de Resampling para complementos verticales

En la siguiente sección de imágenes, figura 5.29, se muestran de izquierda a derecha: la imagen procesada mediante Seam Carving, la imagen original y la imagen procesada mediante Resampling. Este caso en concreto ha sido estudiado previamente en la sección 5.3.1 llegando a la conclusión de que el algoritmo de cambio de relación de aspecto basado en contenido, Seam Carving, ejercía satisfactoriamente su función. En comparación, la imagen con Resampling, aunque no resulta excesivamente molesta, se percibe como incoherente respecto a la original. Ha perdido información puesto que la constitución del niño de casco rojo, que constituye el elemento principal de la imagen, ha cambiado drásticamente.



Figura 5.29: Comparación para imágenes en movimiento

5.3.2.4 Resultados globales

Las medias globales se muestran en la tabla 5.8

	Ratio 0.7	Ratio 0.5
Media total	3.16/5 – 6.32/10	3.61/5 – 7.22/10

Tabla 5.8: Medias totales de ambos sistemas.

En conclusión, para aquellas imágenes que cumplen los requisitos apropiados para una correcta elaboración de la función de energía y desarrollo del algoritmo, la opción de Seam Carving es sin lugar a dudas la elección más adecuada. Para el resto de imágenes, métodos convencionales como el estudiado en esta sección pueden ser una solución suficiente, aunque no garantice las condiciones de calidad que puede llegar a proporcionar Seam Carving.

Una posibilidad a tener en cuenta es que quizás las personas estemos más acostumbradas a las deformaciones propias del resampling, por tratarse de un recurso más habitual en sistemas comerciales y por tanto las imperfecciones que produce no nos llaman tanto la atención como las producidas por un mecanismo menos común como es Seam Carving. Este caso implicaría un sesgo hacia la alternativa de resampling en la evaluación.

6. GESTION DEL PROYECTO

A continuación, a lo largo de este capítulo, se comenta el ciclo de vida del proyecto, los requisitos presupuestarios estimados y las restricciones reguladoras que le afectan.

6.1 ORGANIZACION DE TAREAS

Para la elaboración de este proyecto, se ha seguido un desarrollo en cascada, de modo que para comenzar cada etapa se requiere esperar a la finalización de la etapa anterior, como se muestra en la figura 6.1 [25]. Este desarrollo abarca todo el conjunto de fases comprendidas desde el inicio del proyecto hasta su terminación definitiva. Esto incluye la investigación previa y estudio del problema, la elección justificada de una solución y el planteamiento del diseño, la implementación, las pruebas realizadas, y finalmente la redacción de la documentación referente al proyecto en su totalidad.

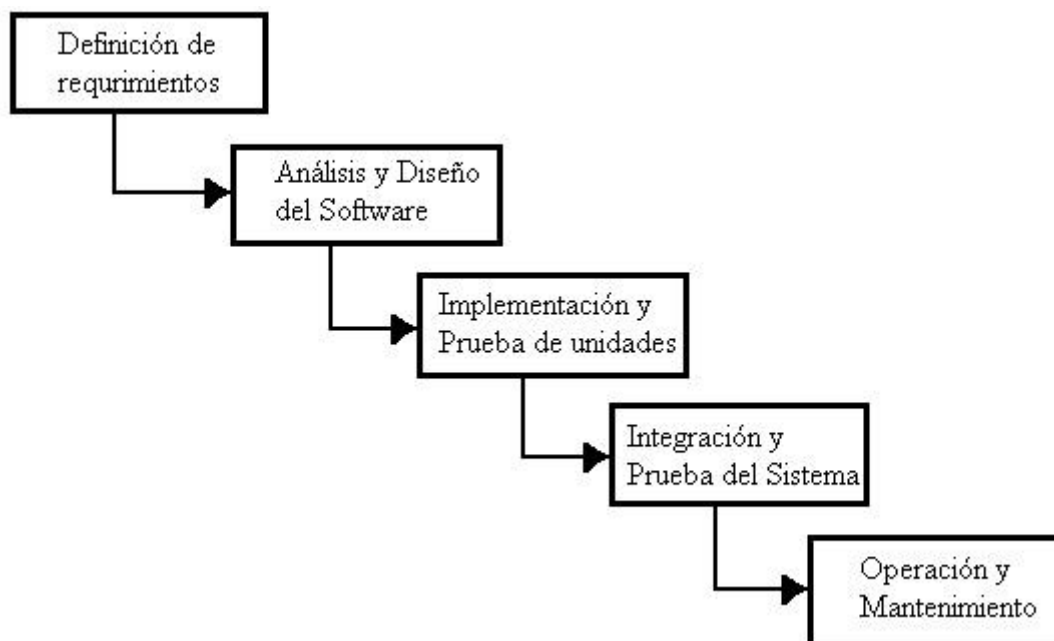


Figura 6.1: Etapas generales del modelo de desarrollo en cascada – Obtenida de http://commons.wikimedia.org/wiki/File:Modelo_Cascada_Secuencial.jpg

A continuación se describe brevemente el desglose de las fases:

Fase 1: Se recoge toda la documentación inicial necesaria y se analizan artículos referentes al método que se pretende estudiar así como otros respecto a métodos alternativos. En este apartado se debe realizar una investigación lo más exhaustiva

posible para tener en cuenta todos los puntos a tratar y contar con una visión global detallada de la situación.

Fase 2: Una vez definido el problema que se desea solventar y las posibles soluciones, el siguiente paso es elaborar el diseño del sistema que va a llevar a cabo dichas soluciones y comenzar la preparación de las herramientas necesarias. En este proyecto se requería confeccionar una base de datos de imágenes con una jerarquía de hasta nivel 2 para poder realizar pruebas posteriores. También hubo que elegir una recomendación para el análisis de resultados y el lenguaje de programación en el que desarrollar el código pertinente teniendo en cuenta los requisitos de software.

Fase 3: La implementación del sistema se divide en la realización de un primer prototipo funcional basado en las definiciones estipuladas durante las fases anteriores y en la mejora y optimización del sistema. Una vez conseguida una funcionalidad adecuada se implementan también los diferentes test que se utilizarán en el proceso de experimentación.

Fase 4: Se llevan a cabo diversas pruebas con el fin de verificar la integridad, usabilidad y eficiencia del programa, aplicando los cambios requeridos si se detectan errores o problemas. Después se comienza con los test previstos para un número estipulado de usuarios y se recogen los resultados para su análisis.

Fase 5: Redacción de la memoria del proyecto, examen en profundidad de los datos recogidos y presentación de los resultados.

En la tabla 6.1 se asocia una cantidad de horas aproximada a cada fase del proyecto.

FASES	HORAS EMPLEADAS
Estudio inicial	25
Elaboración del diseño	25
Desarrollo de la implementación	120
Realización de pruebas	30
Redacción del documento	100
TOTAL	300

Tabla 6.1: desglose de tareas

6.2 PRESUPUESTO

En esta sección se especifica una aproximación de los costes derivados de la realización de las fases señaladas en el apartado anterior.

En la tabla 6.2 se detalla el presupuesto correspondiente al material utilizado, teniendo en cuenta el precio inicial y el coste asociado al proyecto. Principalmente provienen de las prestaciones requeridas para un ordenador de alta gama, para agilizar los procesos computacionales y el procesamiento de la amplia base de datos de imágenes.

Material	Precio	Periodo de amortización	Coste en proyecto
Intel Core i7-3770K 3.5Ghz	230,58€	3 años	19,16€
Gigabyte GeForce GTX 600	165€	3 años	13.75€
Microsoft Windows 7	205€	3 años	17,08€
Microsoft Office 2013	119€	3 años	9,92€
TOTAL	554€	-	62.9€

Tabla 6.2: Costes de material.

A continuación, en la tabla 6.3 se detalla el presupuesto correspondiente a la contratación de personal cualificado para cada tarea.

Puesto	Precio por hora	Horas	Coste total
Jefe de proyecto	90	25	2.250€
Ingeniero	60	275	16.500€
TOTAL	-	300	18.750€

Tabla 6.3: Costes de personal.

Teniendo en cuenta tanto las necesidades materiales como de personal, el coste total de la realización del proyecto viene determinado en la tabla 6.4.

Material	62.9€
Personal	18.750€
COSTE FINAL	18.858,9€

Tabla 6.4: Costes totales.

6.3 MARCO REGULADOR

Seam Carving es un algoritmo muy útil para procesamiento de imagen que puede estar al alcance de cualquier usuario final o empresa, ya sea a través de software libre o de aplicaciones fotográficas de pago que lo incluyan entre sus herramientas, pero es importante resaltar que existen restricciones reguladoras respecto al uso de este tipo de recursos.

La legislación sobre los derechos de autor afecta por supuesto a las imágenes, ya que la propiedad intelectual de una obra literaria, artística o científica corresponde al autor por el solo hecho de su creación. Por tanto, es mandatorio que si hay intención de que las imágenes utilizadas sean reproducidas o distribuidas sea bajo consentimiento de su legítimo autor. Según la Ley de Propiedad Intelectual, artículo 4, se entiende por divulgación de una obra toda expresión de la misma que la haga accesible por primera vez al público en cualquier forma; y por publicación, la divulgación que se realice mediante la puesta a disposición del público de un número de ejemplares de la obra que satisfaga razonablemente sus necesidades estimadas de acuerdo con la naturaleza y finalidad de la misma. [26].

Por otro lado, según el artículo 21 de esa misma ley, la transformación de una obra comprende su traducción, adaptación y cualquier otra modificación en su forma de la que se derive una obra diferente. Los derechos de propiedad intelectual de la obra resultado de la transformación corresponderán al autor de esta última, sin perjuicio del derecho del autor de la obra preexistente de autorizar, durante todo el plazo de protección de sus derechos sobre ésta, la explotación de esos resultados en cualquier forma y en especial mediante su reproducción, distribución, comunicación pública o nueva transformación.

Con este proyecto se busca la capacidad de procesar las imágenes atendiendo al contenido para una optimización, adaptación y modernización de las tecnologías actuales, pero recurrir al uso fraudulento de estos recursos es totalmente ilegal, debe ser evitado y se recomienda a los usuarios tenerlo en cuenta.

Por otro lado, la licencia⁵ de la base de datos de SUN database permite la libertad de copiar y distribuir el material del que disponen en cualquier medio o formato, así como su adaptación que se regirá por los mismos términos que la licencia original. No obstante, deben acreditarse debidamente los términos de la licencia y especificar las modificaciones realizadas en las imágenes utilizadas de esta base de datos. No está permitido, sin embargo, su uso para propósitos comerciales.

⁵ <http://creativecommons.org/licenses/by-nc-sa/3.0/>

7. CONCLUSIONES Y LINEAS FUTURAS

En este apartado se exponen brevemente y de forma sintetizada las conclusiones obtenidas a lo largo de todo el documento, así como las líneas futuras que tratan sobre posibles expansiones del trabajo realizado.

7.1 CONCLUSIONES

En el presente Trabajo Fin de Grado se ha presentado un operador para redimensionamiento de imágenes basado en contenido utilizando el algoritmo *Seam Carving*. Los seams son calculados como los caminos óptimos de píxeles en una misma imagen que cargan la mínima cantidad de energía y son eliminados de ésta consecutivamente hasta obtener la relación de aspecto deseada.

A su vez, y dada la ausencia de un estudio de este tipo en la literatura, se ha llevado a cabo una evaluación sistemática del rendimiento de Seam Carving, atendiendo tanto a las posibles funciones de energía que se emplean, como a la comparación con otras técnicas tradicionales de redimensionamiento de imágenes. Ante la imposibilidad de obtener medidas directas de calidad que correlen adecuadamente con la percepción humana, este proceso de evaluación ha seguido las pautas descritas en la recomendación ITU-R BT.500-11, la cual provee mecanismos que permiten cuantificar las opiniones subjetivas de los observadores.

Esta evaluación se ha combinado con la generación de una base de datos jerarquizada, organizada en categorías y subcategorías relacionadas con el contenido mostrado en cada imagen.

En cuanto al análisis de resultados, se ha concluido que el cálculo de la función de energía utilizada para establecer los seams es tan bueno, e incluso ligeramente mejor, si se realiza a través de la norma 1 de los gradientes de la imagen que utilizando técnicas más avanzadas como los modelos de saliencia visual.

A su vez, se ha observado cómo este tipo de aproximación supera en rendimiento a una técnica tradicional del redimensionamiento de imágenes como es el remuestreo.

Por último, analizando los resultados particulares para cada una de las categorías de imágenes, la técnica de Seam Carving ha demostrado obtener resultados gratamente sorprendentes para muchas categorías de imágenes, principalmente en aquellas

escenas de tipo natural en la que no aparecen elementos creados por el hombre. No obstante, su rendimiento aún puede perfeccionarse para muchos otros tipos de imágenes basándose en los defectos y limitaciones encontrados durante las pruebas.

7.2 LINEAS FUTURAS

Existen numerosas posibilidades de extensión de este trabajo. Una de ellas es extender esta aproximación hacia otros dominios, el primero de ellos el redimensionamiento en vídeo [29]. Esto implicaría pasar de seams exclusivamente espaciales a seams espacio-temporales [30]. Otra es la combinación de diferentes métodos para perfeccionar los resultados, puesto que diferentes métodos actúan con mayor o menor eficiencia dependiendo de las características de la imagen. Una implementación básica sería compaginar adecuadamente la eliminación de seams verticales y horizontales simultáneamente, pudiendo tener un orden que priorizase los seams de forma entremezclada para optimizar los resultados.

Este operador, Seam Carving, podría utilizarse también para tareas de amplificación de contenido, no sólo reducción, y supresión de elementos en una imagen. Además, puede ser fácilmente integrado con varios tipos de medidas de saliencia y con parámetros de entrada seleccionables por el usuario, para guiar el proceso de redimensionamiento.

8. REFERENCIAS Y BIBLIOGRAFIA

- [1] Shai Avidan and Ariel Shamir, "Seam Carving for Content-Aware Image Resizing", Journal ACM Transactions on Graphics (TOG), Proceedings of ACM SIGGRAPH 2007, vol. 26, Issue 3, article No. 10, July 2007.
- [2] Haipeng Jia, Yunguan Zhang, Weiyan Wang and Juanliang Xu, "Accelerating Viola-Jones Face Detection Algorithm on GPUs", HPCC IEEE 9th International Conference on Embedded Software and Systems, Pages 396-403, 2012
- [3] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No 11, pp. 1254-1259, Nov 1998.
- [4] Jorge Calderón, CEO de U-tad y miembro del Cluster ICT-Audiovisual de Madrid, "Estado del arte de las tecnologías audiovisuales" 3ª edición, nota de prensa, 2013.
- [5] Matt Ward, "One Unexplored Reason Why Photoshop Won't Be Dethroned Anytime Soon", Graphic Design, <http://www.myinkblog.com/>
- [6] Adobe Photoshop, Adobe Community Help, "Escala según el contenido", <http://helpx.adobe.com/es/photoshop/using/content-aware-scaling.html>
- [8] D. Fernandez-Duque and M.I. Posner (2001) "Brain imaging of attentional networks in normal and pathological states". Journal of Clinical and Experimental Neuropsychology, 57, pages 74–93. DOI: 10.3758/bf03206794.
- [9] A. M Treisman and Gelade (1980) "A feature-integration theory of attention". Cognitive psychology 12: 97–136. DOI: 10.1016/0010-0285(80)90005-5.
- [10] Dirk Bernhardt Walther, Caltech student final project "Interactions of visual attention and object recognition: computational modeling, algorithms, and psychophysics." (2006) Chapter 2: A Model of Salient Region Detection. Caltech Library.
- [11] J. Duncan and G.W. Humphreys (1989) "Visual search and stimulus similarity" Psychological Review 96, pages 433-458. DOI: 10.1037/0033-295x.96.3.433.
- [12] Backus, John W. Can Programming Be Liberated from the von Neumann Style? A Functional Style and Its Algebra of Programs. DOI: 10.1145/359576.359579.
- [13] Simons & Levin (1997) and O'Regan, Rensink & Clark (1999) "Change Blindness".

- [14] Laurent Itti, Christof Koch. "A saliency-based search mechanism for overt and covert shifts of visual attention" (1999).
- [15] Ali Borji and Laurent Itti. - State-of-the-Art in Visual Attention Modeling - IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 35, No. 1, JANUARY 2013.
- [16] L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Trans. Pattern Analysis (Nov. 1998).
- [17] D. Walther and C. Koch, "Modeling Attention to Salient Proto-Objects," Neural Networks, vol. 19, no. 9, pp. 1395-1407, 2006.
- [18] S. Frintrop, VOCUS: A Visual Attention System for Object Detection and Goal-Directed Search. Springer, 2006.
- [19] J. Harel, C. Koch, and P. Perona, "Graph-Based Visual Saliency," Neural Information Processing Systems, vol. 19, pp. 545-552, 2006.
- [20] Mathworks - <http://www.mathworks.es/products/matlab/>
- [21] Graph Based Visual Saliency - www.vision.caltech.edu/~harel/share/gbvs/
- [22] SUN DATABASE - <http://groups.csail.mit.edu/vision/SUN/>
- [23] RECOMMENDATION ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures.
- [24] Scott Sehlhorst, Software Process (Waterfall Process versus Incremental Process), January 3rd, 2006, and D. Herbert Benington, "Production of Large Computer Programs".IEEE Annals of the History of Computing, 1983. IEEE Educational Activities Department 5 (4): 350–361. DOI: 10.1109/MAHC.1983.10102. Retrieved 2011-03-21.
- [25] Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia.
- [27] Antonio Torralba and Aude Oliva, Massachusetts Institute of Technology, "Contextual Guidance of Eye Movements and Attention in Real-World Scenes: The Role of Global Features in Object Search" Psychological Review, American Psychological Association 2006, Vol. 113, No. 4, 766–786.

- [28] T. Judd, K. Ehinger, F. Durand, and A. Torralba. "Learning to predict where humans look" IEEE International Conference on Computer Vision (ICCV), 2009.
- [29] M. Decombas, F. Dufaux, B. Pesquet-Popescu, "Spatio-temporal grouping with constraint for seam carving in video summary application", Digital Signal Processing (DSP), 2013 18th International Conference on, pages 1 – 8.
- [30] J. Rügge, O. Wang, A. Smolic, M. Gross, "DuctTake: Spatiotemporal Video Compositing", Eurographics 2013, Volume 32, Number 2. http://zurich.disneyresearch.com/~owang/pub/pdfs/ducttake_authorcopy.pdf
- [31] Arne Nowak (October 2010). "Digital Cinema Technologies from the Archive's Perspective". p. 4. Retrieved August 8, 2013.
- [32] Louis Goddard, "Philips discontinuing super-wide Cinema 21:9 TVs due to lack of demand". The Verge. 2012-08-28.
- [33] Michael Elad – "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing" Springer, 2010.
- [34] Cloonan, M.V. and Sanett, S. "The Preservation of Digital Content," Libraries and the Academy. Vol. 5, No. 2 (2005): 213-37.
- [35] Krista A. Ehinger and Barbara Hidalgo-Sotelo, Antonio Torralba and Aude Oliva , "Modelling search for people in 900 scenes: A combined source model of eye guidance", Psychology Press, Visual Cognition, 2009, 17 (6/7), 945-978.
- [36] Raanan Fattal, Dani Lischinski, Michael Werman, "Gradient Domain High Dynamic Range Compression", School of Computer Science and Engineering, ACM SIGGRAPH 2002 <http://www.cs.huji.ac.il/~danix/hdr/hdrc.pdf>
- [37] Azriel Rosenfeld, Picture Processing by Computer, New York: Academic Press, 1969.

